
United States
Department of
Agriculture
National Agricultural Library

Final Report
September 1998
On Site Interactive Model for Irrigation Management

Dale F. Heermann, Richard Davis, Sarah Streett,
La Verne Stetson, Steffen Meyer
USDA-ARS-NPA

1 Introduction

Irrigation pumping is concentrated during a three- to five-month period for much of the pump-irrigated areas of the U.S. Electric powered pumping increases the peak demand of the electric utilities serving these irrigation loads whether they are wholesale or retail loads. Although pumping increases peak demands for the electric power supply system, it is also a significant part of the annual revenue for some utilities and contributes to the economies of communities in irrigated areas.

1.1 Objective

Our objective was to evaluate or develop electrical load prediction programs and irrigation scheduling programs that predict irrigation requirements for integration into an interactive load management program that could be used by electric utilities for controlling peak demands where an irrigation load is a significant part of the demand.

1.2 Procedure

The procedure was to review the literature, evaluate the AGEND Irrigation Model for forecasting irrigation electrical use, establish and meet with a customer focus group and acquire data to begin development of a forecasting model.

Many of the attached references were well known to the principal investigators. They were used for the AGEND model and several studies completed under the guidance of the principal investigator. The later studies were conducted in the 1970's and studied the load management from the irrigator's viewpoint. The current study was focused at the level of the supplier to the irrigator. The forecasting model that was evaluated and developed will compliment previous studies. A number of references identified in the automated literature search were written by the principal investigators and are cited here.

The AGEND Irrigation Model was developed for EPRI under a contract with Quantum Consulting, Inc. AGEND is a detailed model, but when given the necessary input data it should provide a good forecast of electrical demand. The model has not been readily accepted by the utilities. The discussion with potential users with the customer focus group showed that the model required more detailed data than is readily available. The model required that the crop must be known for each field irrigated. In addition, the data required for each system included pump efficiency, motor size, pumping lift, pipe sizes and discharge. The estimated running time is generated from crop water requirements. Though the model can give accurate electrical usage, obtaining the detailed information necessary to run the model is difficult.

A customer focus group made up of TRI-STATE G&T technical representatives and most of the REA rural managers and technical specialists with significant irrigation loads in eastern Colorado

and western Nebraska provided input on their needs for forecasting electrical loads. They indicated the type of data that are available for use in a model. Several cooperatives are either considering or are starting load control programs for irrigation loads. They stated that their need is for short term forecasting of irrigation loads with emphasis on the shoulder months of June and September. Just forecasting the next couple of days would be valuable for use in load control programs. This would allow control to be minimized and provide the most benefit for their program.

They identified the types of data that are readily available through either Coop offices or from TRI-STATE. The available data varied from Coop to Coop but usually included billing point demand and total kWh delivered. Sometimes the crop, demand, kWh, location, water use and type of irrigation system are available. The customers were all willing to share data for our analysis. We have obtained the demand data from all billiing points from Highline Electric for our first analysis. The service areas also have weather station data available which will allow calculation of crop water use.

Our initial approach was to perform a time series analysis of the demand data for the years 1992-1996. The S-Plus statistical package was purchased for doing this part of the analysis. The objective was to first determine the maximum information that could be obtained from the data that was readily available and was already computer compatible. The data were obtained directly from TRI-STATE which had edited the data for quality assurance.

We outlined our approach to this study to TRI-STATE personnel and obtained their ideas as to the needs and benefits for load management. We need to understand not only the needs of the rural cooperatives and their constraints but also the objectives and approach of the supplier to deliver the necessary power to meet demand. It was recognized that though each has the objective of reducing costs, the paths to solution can often lead to conflict. It confirms the benefits for the rural cooperatives for a better procedure to forecast short term demands. The electrical delivery system is one where each entity is optimizing within the constraints of their own system and the operating procedures of the other.

The demand data from Highline Electric were obtained for the years 1992-1996. Climatic data from the area were used to estimate crop water use and rainfall amounts. The initial analysis was to develop correlations between the daily peak demands and the climatic data. The primary correlation was determined between the rainfall and sharp drop in peak demands from one day to the next. As interesting as this may be, it did not provide the needed estimate of the future peak daily demands. The relationship between peak demands and ETA was not found to be significant. It may be that the irrigators did not schedule irrigations based on crop ET but based their decisions on perception and historical experience for scheduling irrigations.

The remainder of this report discusses our efforts to create an algorithm for predicting the maximum daily power usage which could be implemented with ease. It was desired that this algorithm be capable of making predictions several days in advance. In particular, seven-step predictions were of interest in order to allow the involved parties sufficient time to react.

The data analyzed in detail was the power demand data (in kilowatts) recorded in thirty minute intervals at the Holyoke station for the months of April through October, years 1992 to 1996. Additionally, daily weather data in the general vicinity of the Holyoke station for the months of April through September were included in the analysis. The climatic data consisted of maximum daily temperature, minimum daily temperature, daily precipitation (in mm), ETP (as calculated by the Penman equation), and ETAcorn (a variation of ETP adjusted for corn, the major crop in the area).

2 Exploratory Plots

2.1 Condensing the Original Power Usage Data

The Holyoke power data were first condensed by finding the maximum power usage for each day and discarding the remaining points. Since our goal was to predict maximum daily power usage, no information seemed to be lost in using these subsets of daily maxima as our working data sets. This was apparent when comparing time series plots of the original data and the daily maximum data subset depicted in Figures 1 through 10. Additionally, we only considered the power data through the month of September since most irrigations are completed by mid September.

2.2 Data Exploration

We examined several time series plots and scatter plots in order to obtain a better understanding of how peak usage changed over time and what were some of the factors influencing this movement. Initially, we considered the relative differenced data X_t ($X_t = \frac{Y_{t+1}-Y_t}{Y_t}$, where Y_t is the maximum daily power usage) in an effort to create a more stationary sequence. This series is depicted for all available years in Figures 11 through 15. The darkened circles indicate that the precipitation on day t was positive ($R_t > 0mm$). In studying these graphs, it appears that there is an inverse relationship between rainfall and power usage. This relationship is depicted more clearly in the scatter plots of X_t vs. R_t (Figures 16 through 20). A regression line was added to these plots for clarification. Although the strength of the relationship between these two variable changes from one year to the next, it is evident in all five years.

Scatter plots involving the remaining weather variables did not reveal any further relationships. In addition to our consideration of various scatter plots, we also examined several brush plots in the hopes of uncovering further associations among the variables. A brush plot is an S-Plus graphing device that allows the user to interact with various 2-dimensional and 3-dimensional scatter plots. One of the options allows the user to highlight specific points on one graph resulting in these points being highlighted on all of the remaining graphs. This is often useful for discovering relationships which are only present for a subset of the data. The user may also change the variable axes on the 3-dimensional graph to explore relationships between any three variables and rotate this graph to obtain different perspectives. However, no further relationships were discovered using this procedure. Our next step was to begin modeling the series.

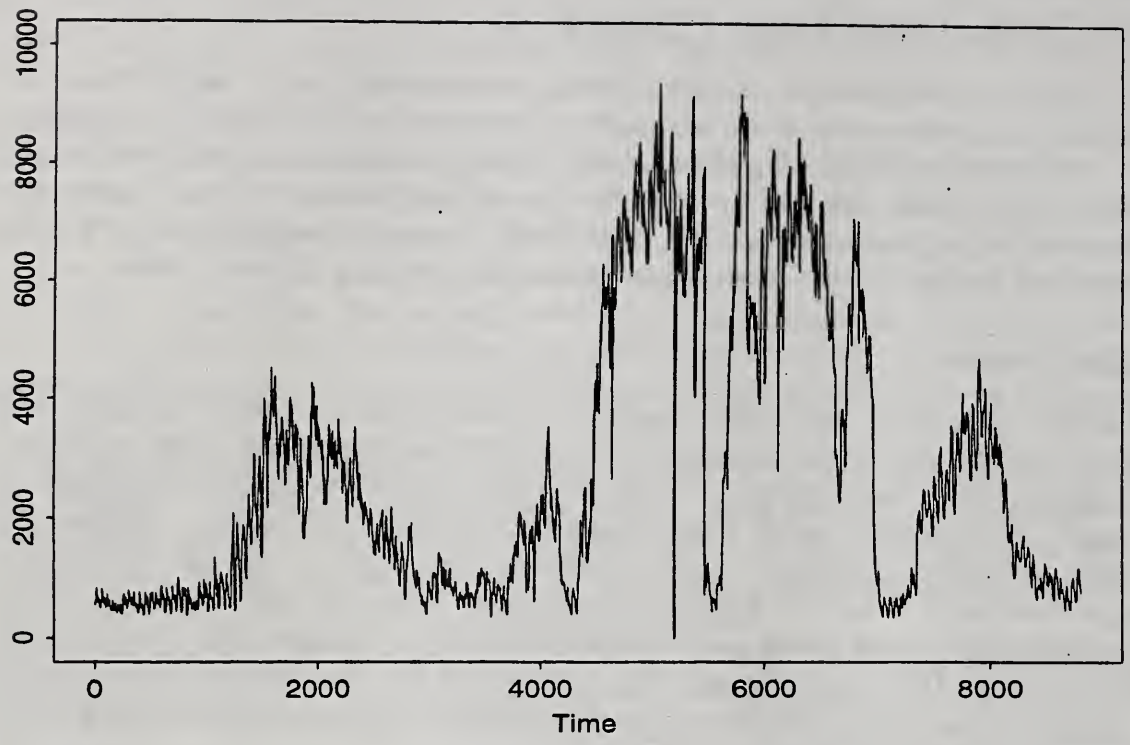

Figure 1: Time Series Plot of Power Usage Data for 1992

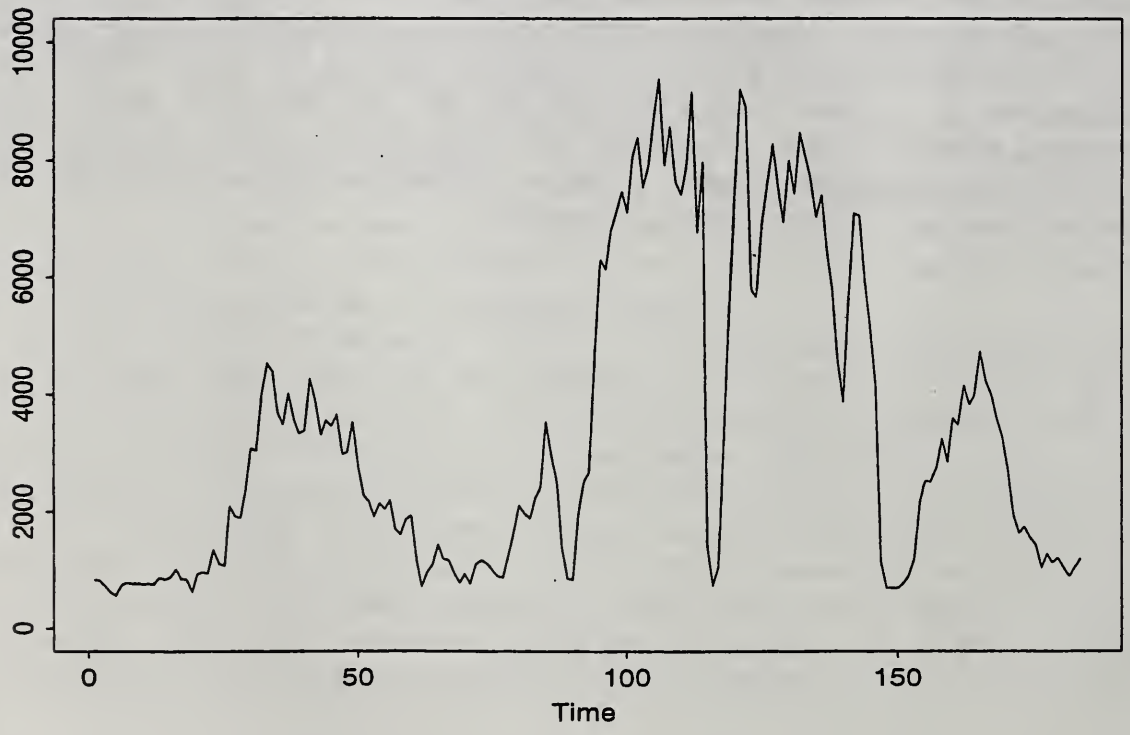

Figure 2: Time Series Plot of Maximum Daily Power Usage for 1992

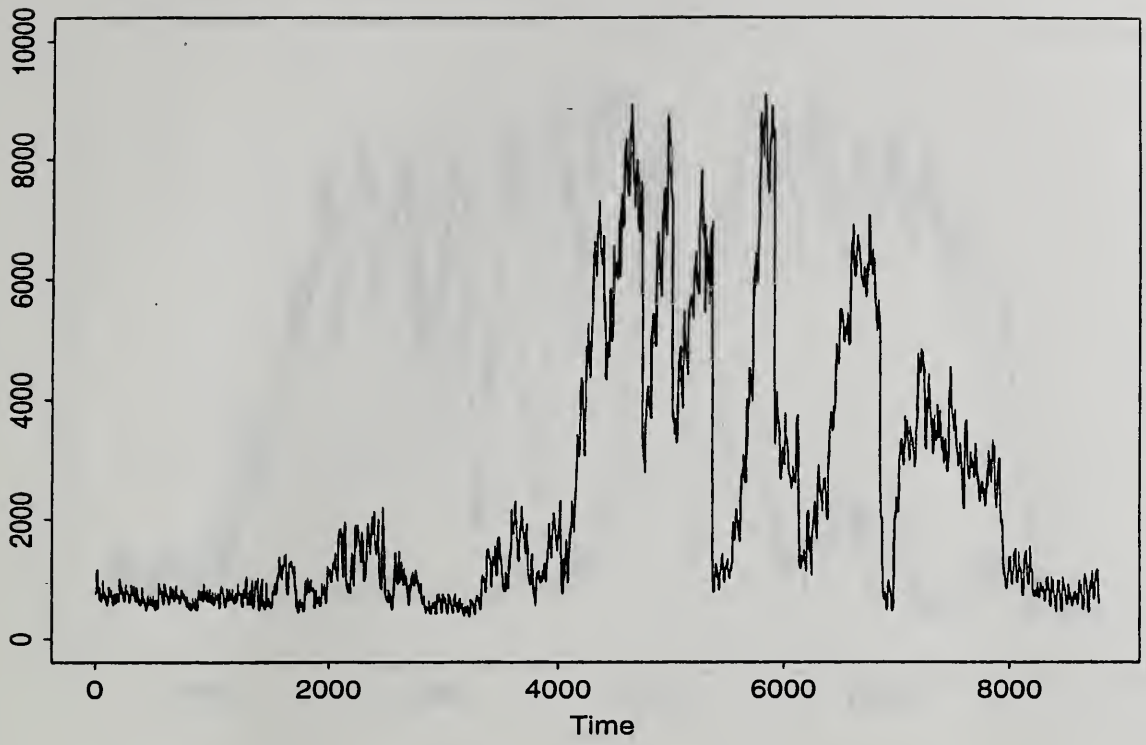

Figure 3: Time Series Plot of Power Usage Data for 1993

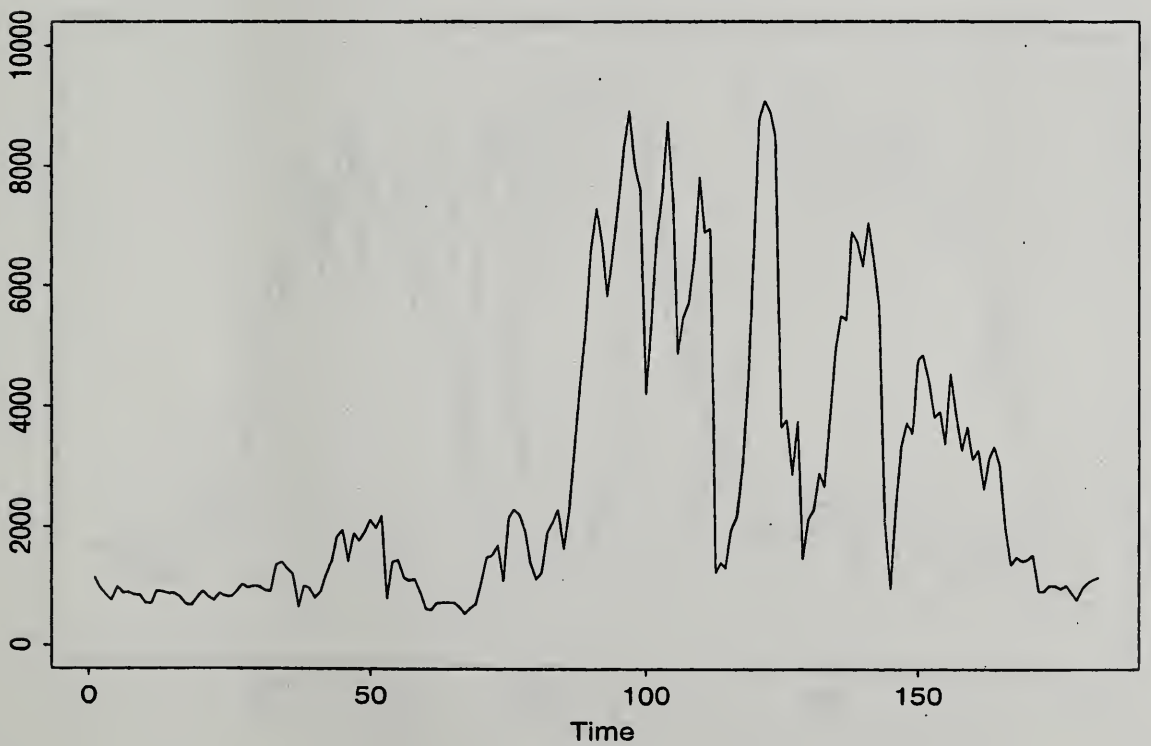

Figure 4: Time Series Plot of Maximum Daily Power Usage for 1993

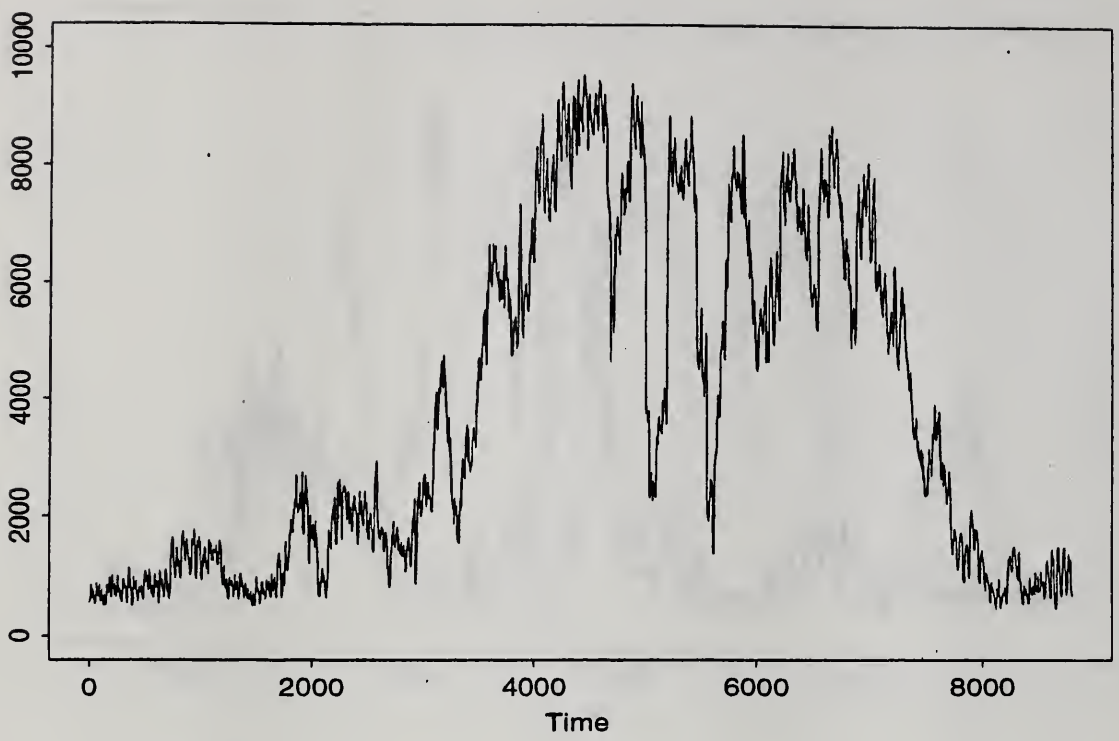

Figure 5: Time Series Plot of Power Usage Data for 1994

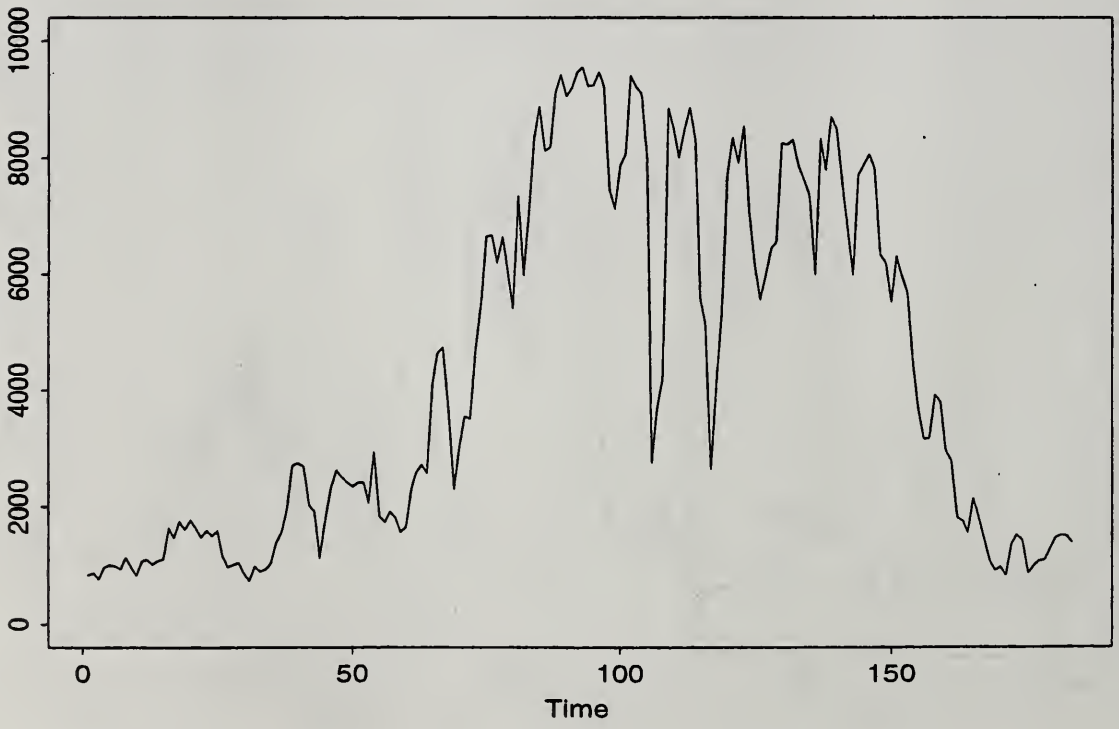

Figure 6: Time Series Plot of Maximum Daily Power Usage for 1994

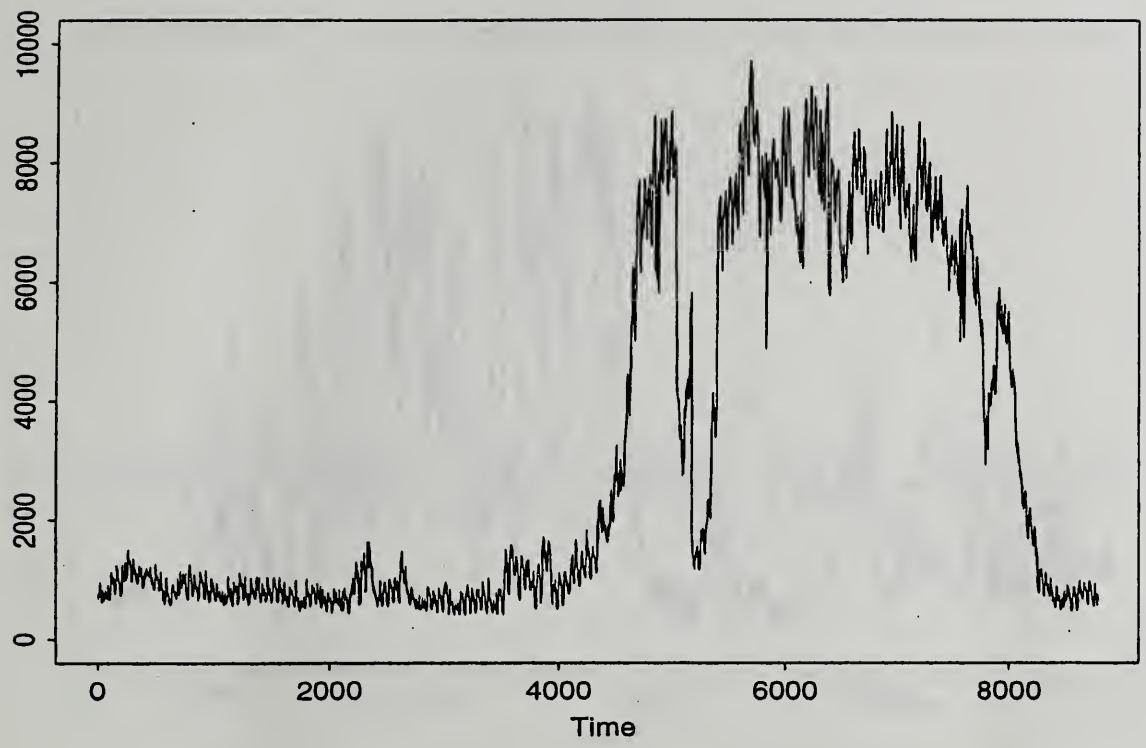

Figure 7: Time Series Plot of Power Usage Data for 1995

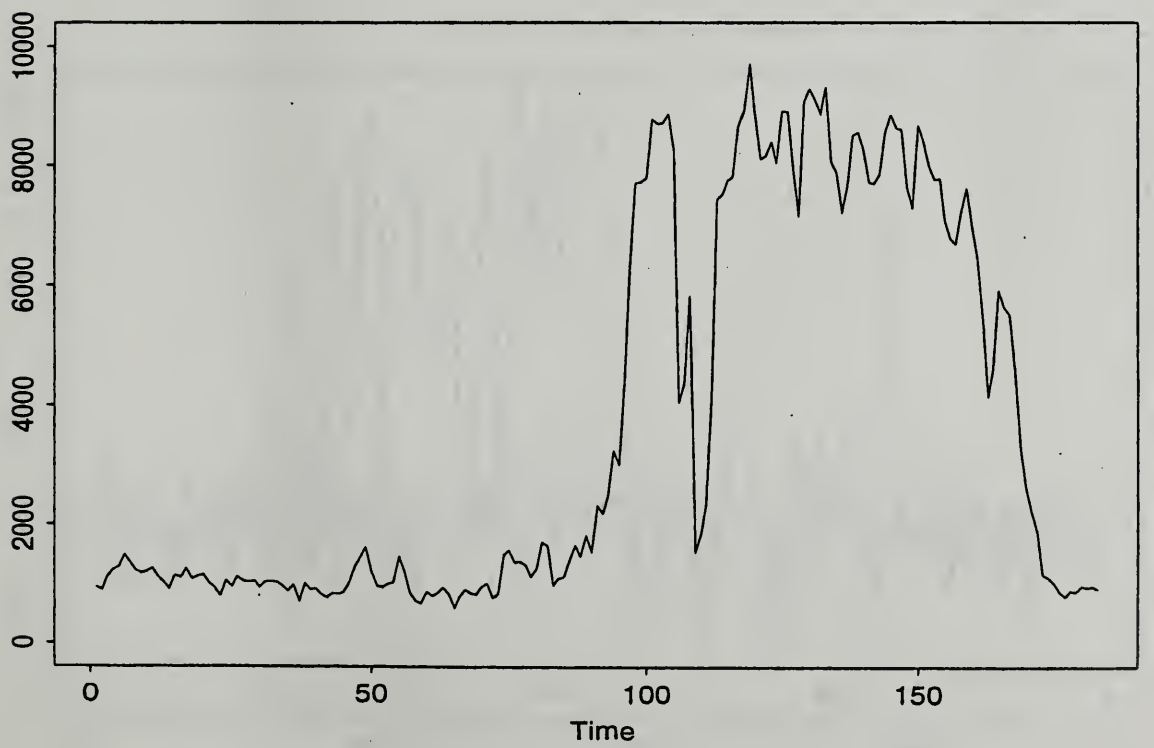

Figure 8: Time Series Plot of Maximum Daily Power Usage for 1995

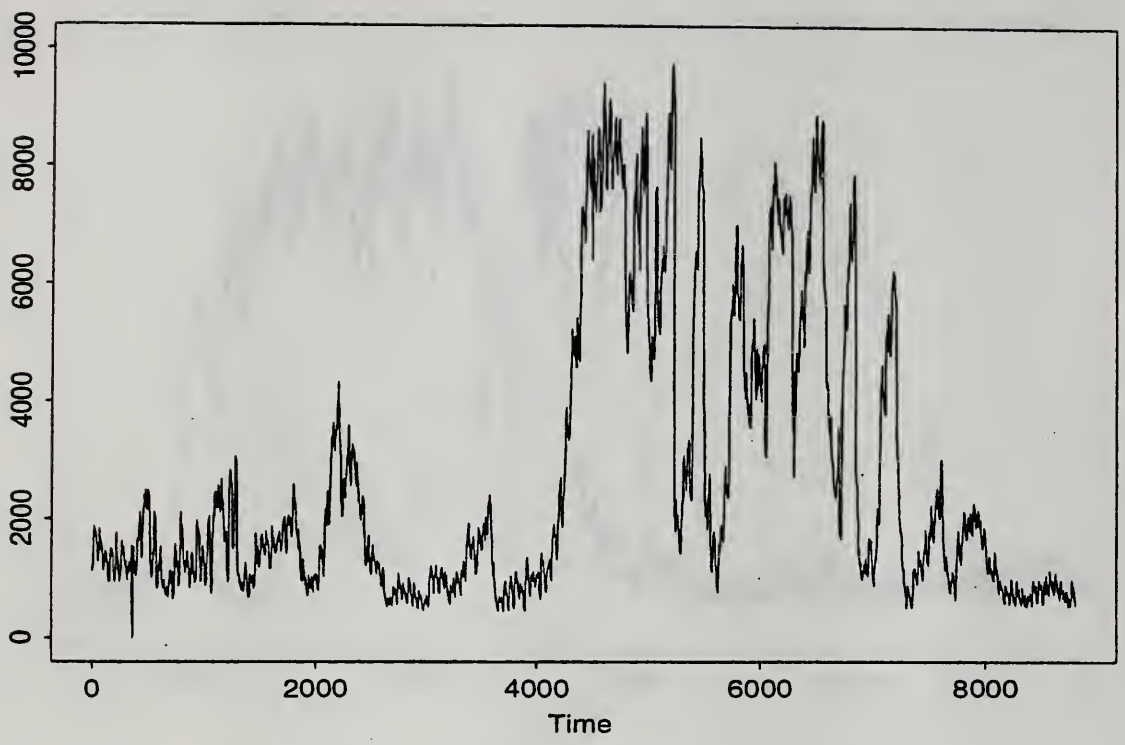

Figure 9: Time Series Plot of Power Usage Data for 1996

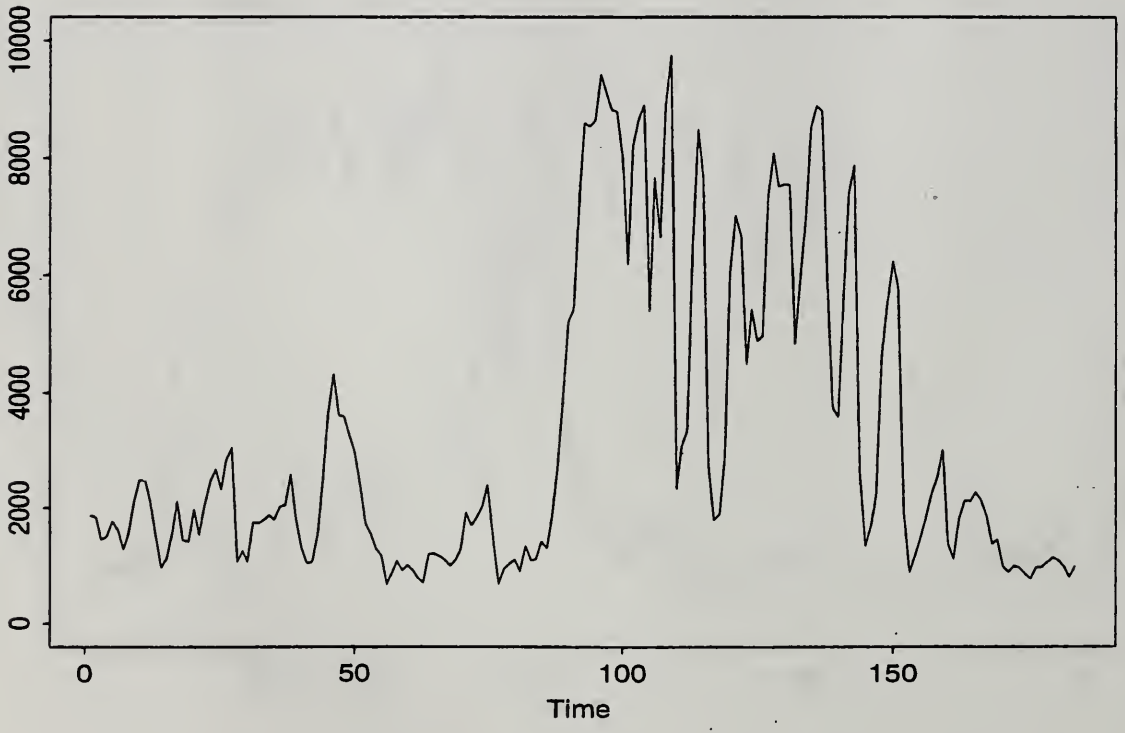

Figure 10: Time Series Plot of Maximum Daily Power Usage for 1996

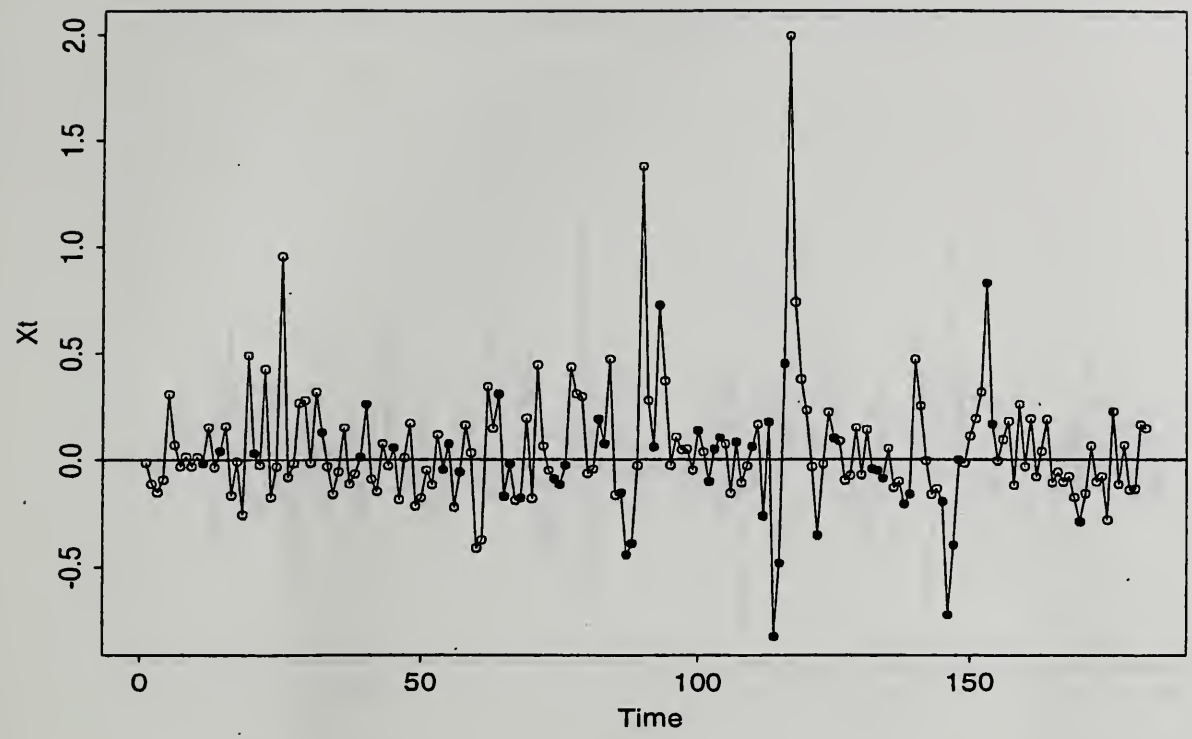

Figure 11: Time Series Plot of Relative Differenced Data for 1992

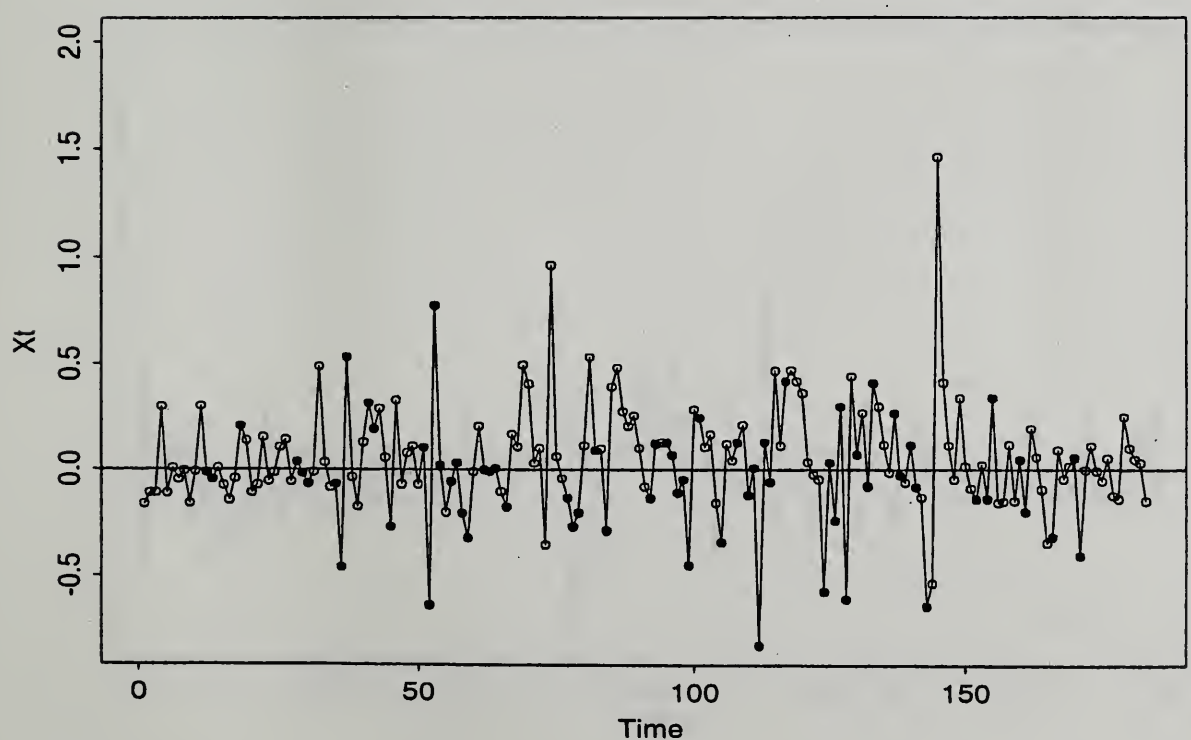

Figure 12: Time Series Plot of Relative Differenced Data for 1993

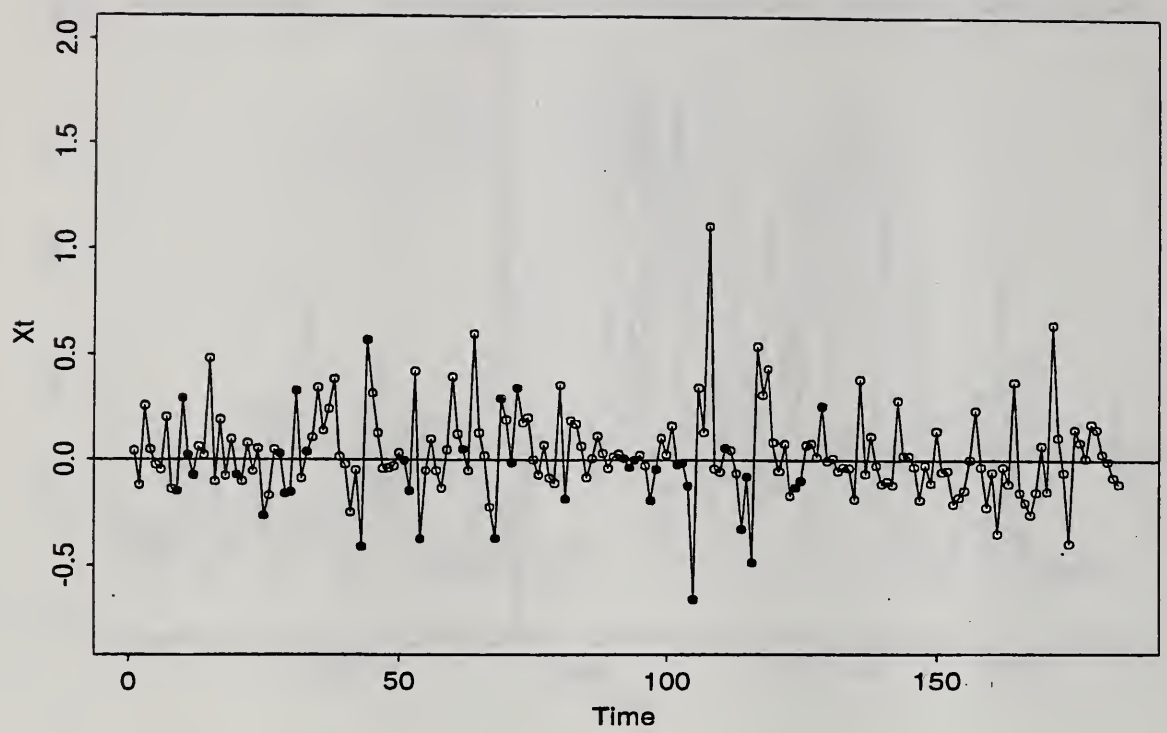

Figure 13: Time Series Plot of Relative Differenced Data for 1994

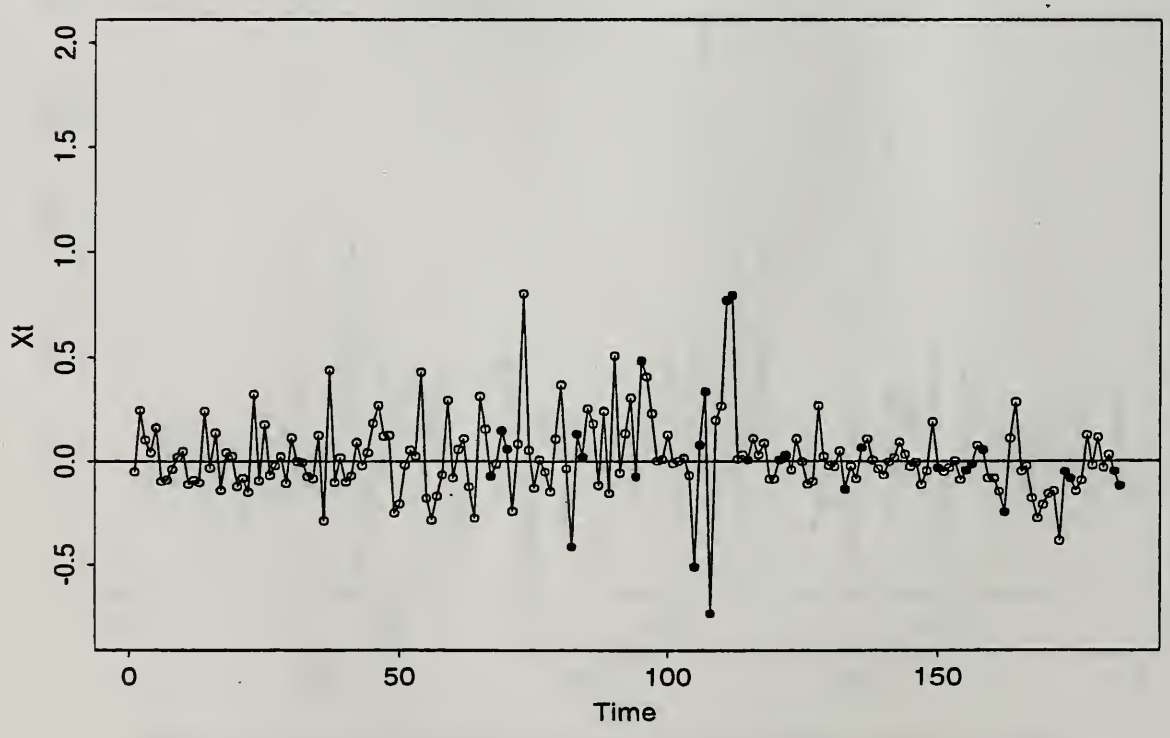

Figure 14: Time Series Plot of Relative Differenced Data for 1995

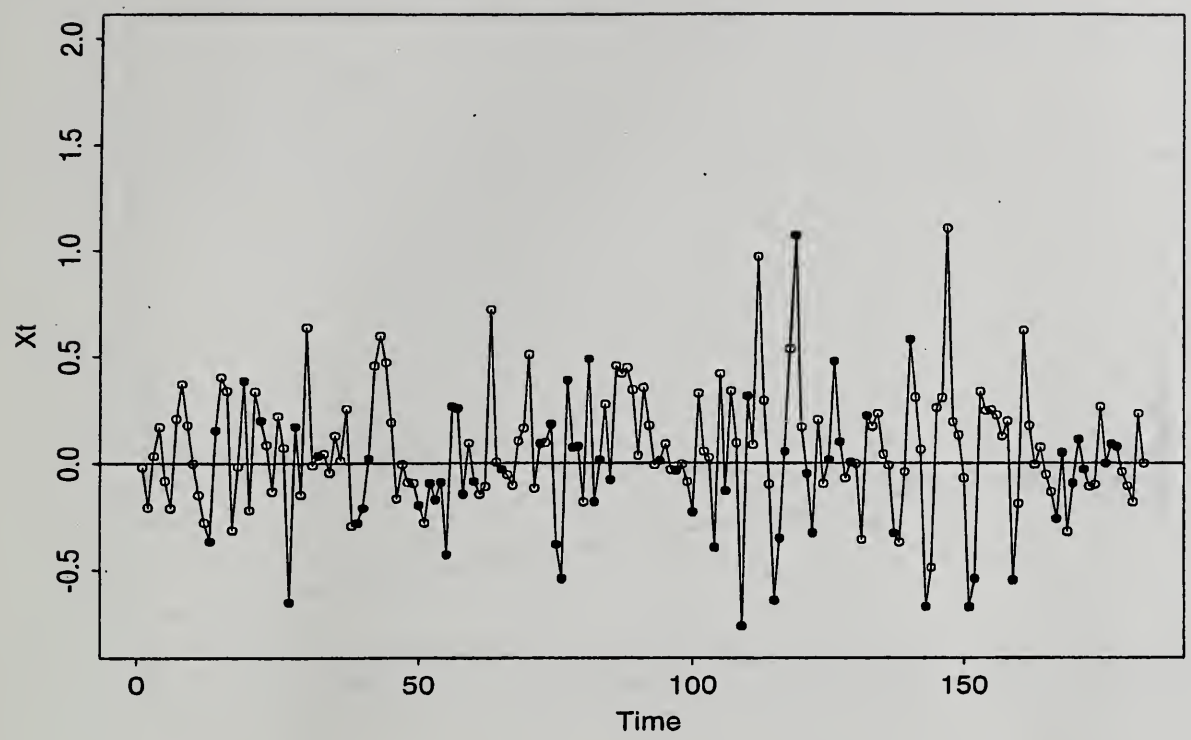

Figure 15: Time Series Plot of Relative Differenced Data for 1996

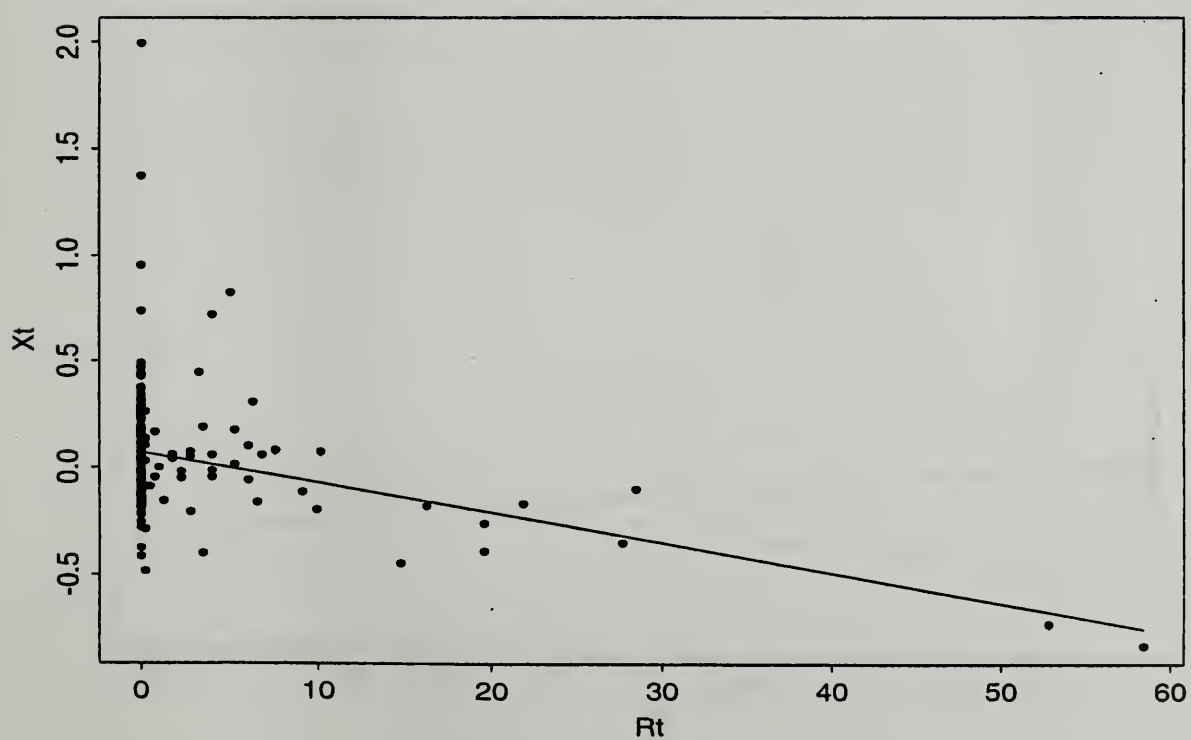

Figure 16: Scatter Plot of X_t versus R_t for 1992

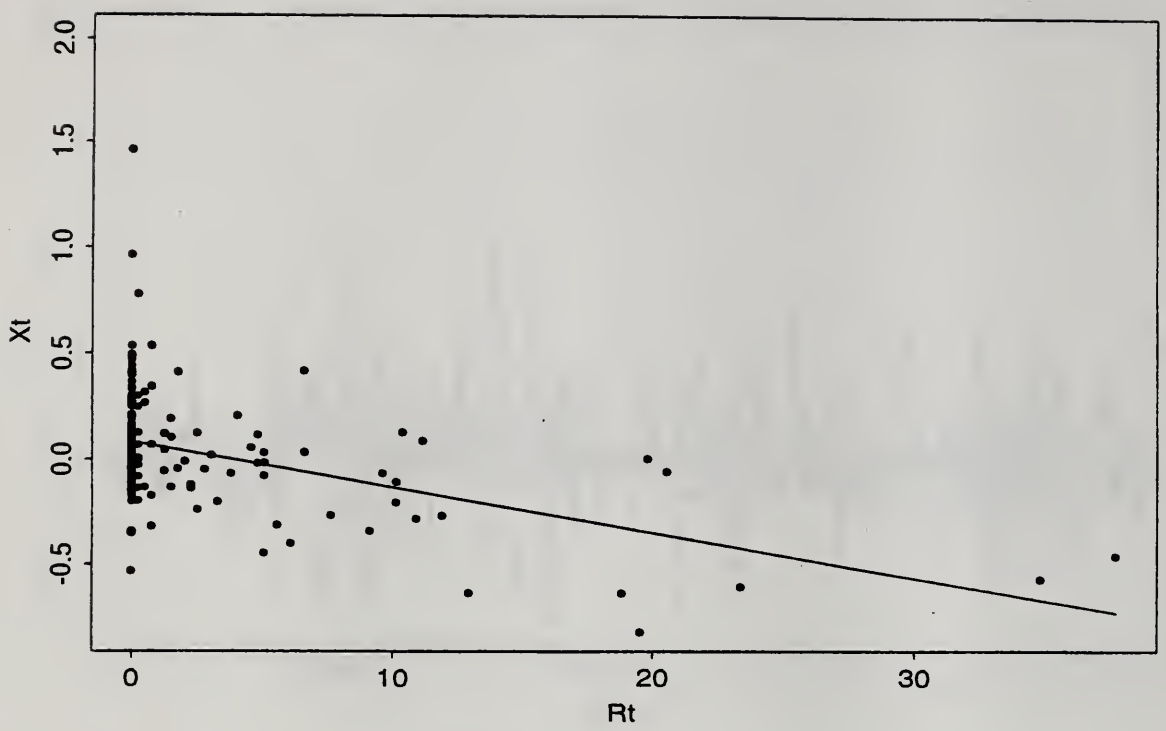

Figure 17: Scatter Plot of X_t versus R_t for 1993

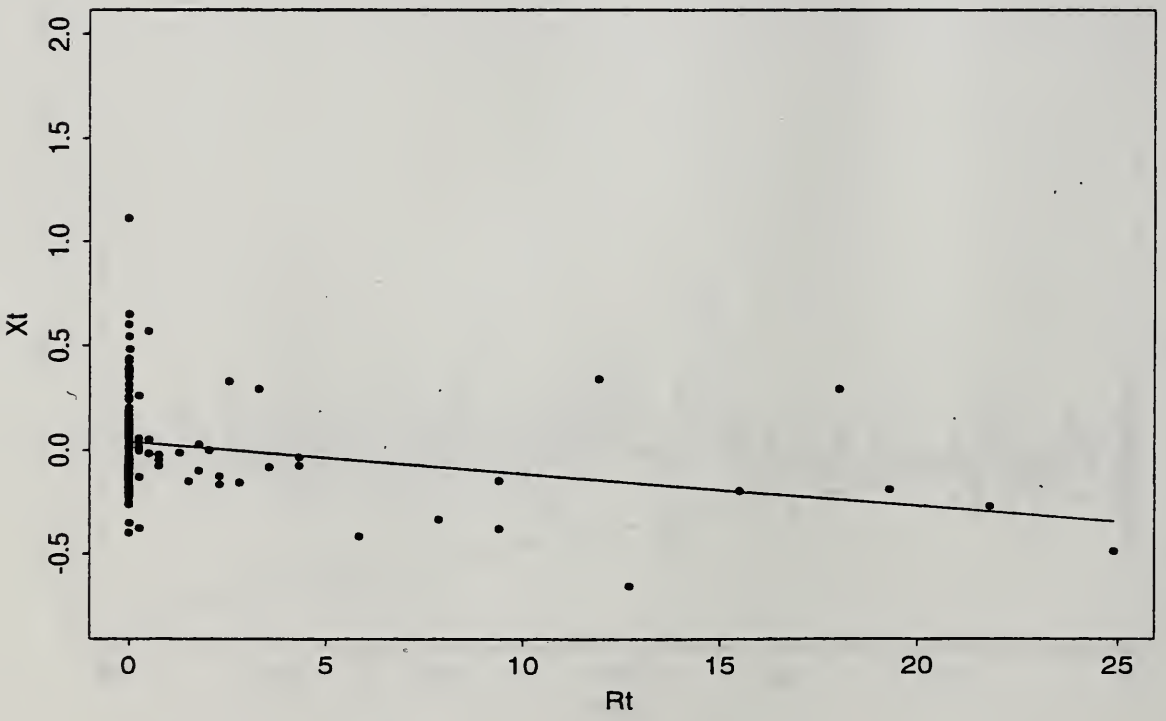

Figure 18: Scatter Plot of X_t versus R_t for 1994

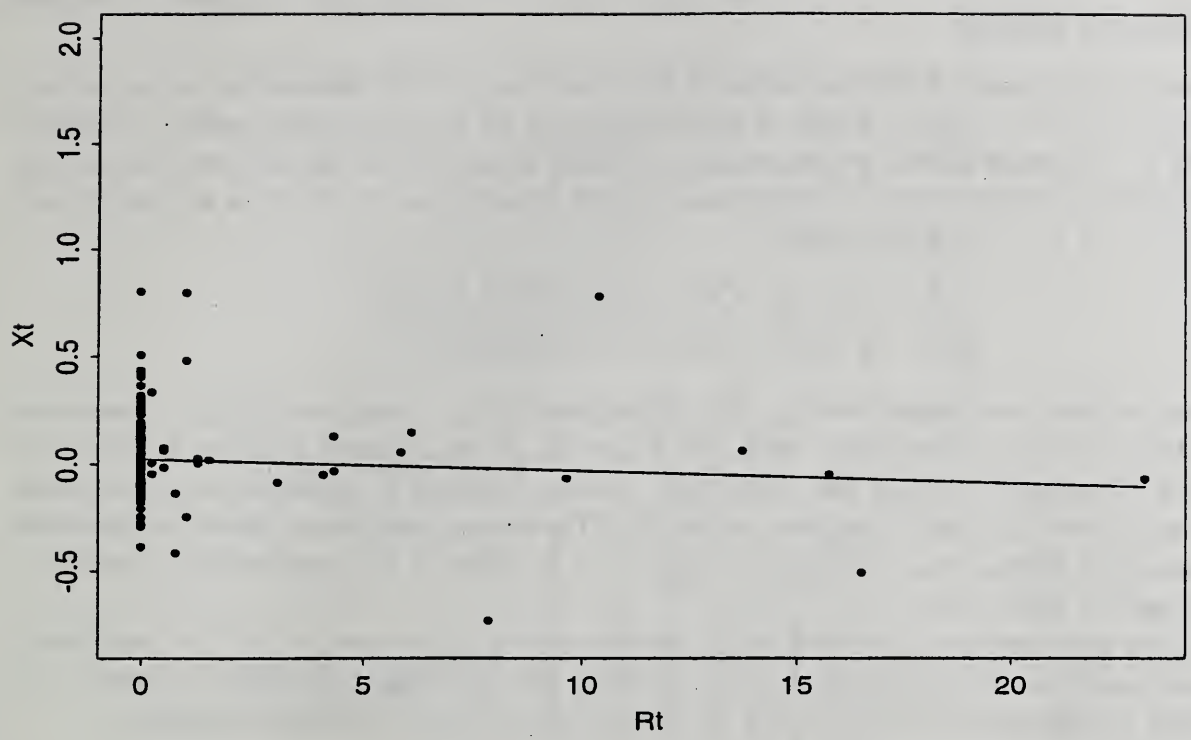

Figure 19: Scatter Plot of X_t versus R_t for 1995

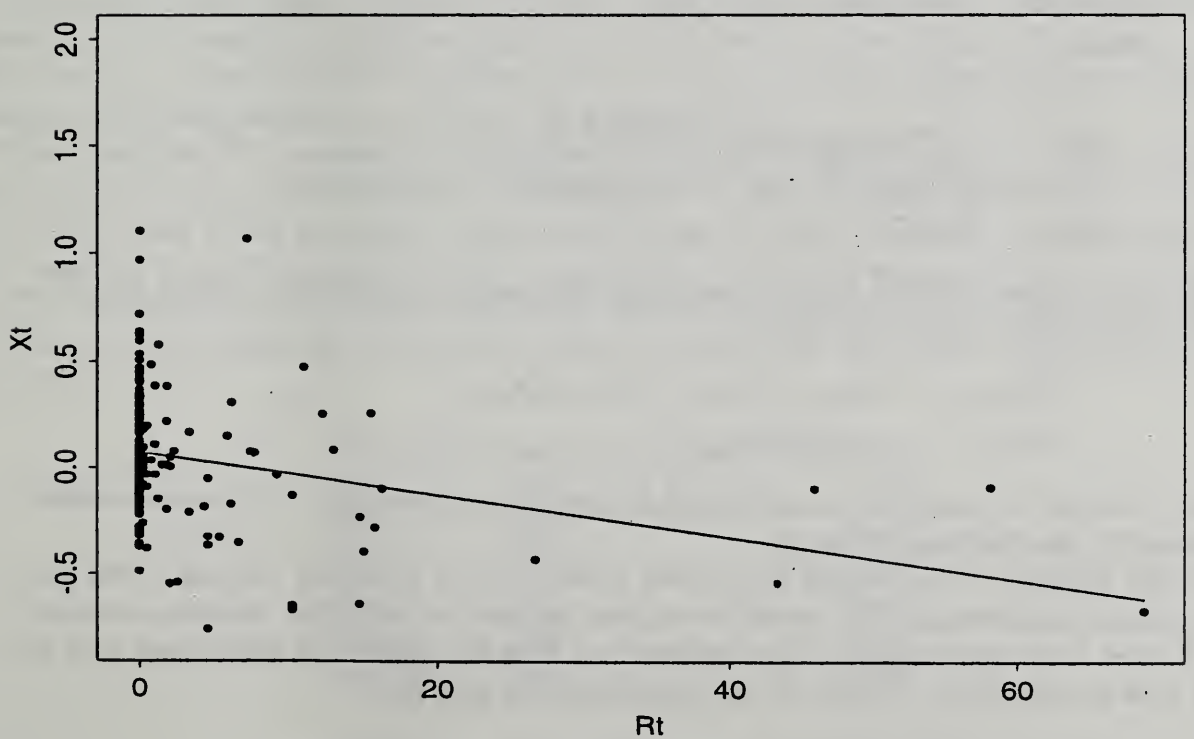

Figure 20: Scatter Plot of X_t versus R_t for 1996

3 Initial Model Considerations

3.1 Structural Models

In order to model the relative differenced series we began with a structural time series model written in state space form. This type of model is often thought to be a good starting point because it encompasses a wide range of time series models, and with the aid of the Kalman recursions, the analysis is relatively straightforward. The general form of a state space model for a w-dimensional time series $\{\mathbf{Y}_t, t = 1, 2, \ldots\}$ is given below.

$$\begin{aligned} \mathbf{Y}_t &= G_t\mathbf{X}_t + \mathbf{W}_t, \ \mathbf{W}_t \sim \mathrm{WN}(\mathbf{0}, \{R_t\}) \\ \mathbf{X}_{t+1} &= F_t\mathbf{X}_t + \mathbf{V}_t, \ \mathbf{V}_t \sim \mathrm{WN}(\mathbf{0}, \{Q_t\}) \end{aligned}$$

where $\mathbf{X}_t$ is a v-dimensional state variable, $\{G_t\}$ is a sequence of $w \times v$ matrices, $\{F_t\}$ is a sequence of $v \times v$ matrices, $\{\mathbf{V}_t\}$ is uncorrelated with $\{\mathbf{W}_t\}$ and $\mathbf{X}_1$ is uncorrelated with both $\{\mathbf{V}_t\}$ and $\{\mathbf{W}_t\}$. The first equation is called the observation equation because it expresses the observation vector, $\mathbf{Y}_t$, as a linear function of the state vector, $\mathbf{X}_t$. The second equation is known as the state equation because it defines the state vector at time $t+1$ in terms of the previous state vector at time t along with a noise term.

The Kalman equations are a recursive set of formulas which can be used to find the best linear mean-squared predictor of the state vector, $\mathbf{X}_t$, based on the information available at time $t-1$. This predictor is denoted by $\hat{\mathbf{X}}_t$ or $P_{t-1}(\mathbf{X}_t)$ and is determined by the following equations:

$$\begin{aligned} \hat{\mathbf{X}}_1 &= P(\mathbf{X}_1|\mathbf{Y}_0) \\ \Omega_1 &= E[(\mathbf{X}_1 - \hat{\mathbf{X}}_1)(\mathbf{X}_1 - \hat{\mathbf{X}}_1)'] \\ \hat{\mathbf{X}}_{t+1} &= F_t\hat{\mathbf{X}}_t + \Theta_t\Delta_t^{-1}(\mathbf{Y}_t - G_t\hat{\mathbf{X}}_t) \\ \Omega_{t+1} &= F_t\Omega_t F_t' + Q_t - \Theta_t\Delta_t^{-1}\Theta_t', \end{aligned}$$

where

$$\begin{aligned} t &= 1, 2, \ldots, \\ \Omega_t &= \text{the error covariance matrix of } \hat{\mathbf{X}}_t \\ \Delta_t &= G_t\Omega_t G_t' + R_t, \ \Delta_t^{-1} \text{ is any generalized inverse of } \Delta_t \\ \Theta_t &= F_t\Omega_t G_t'. \end{aligned}$$

It also follows from these equations that the best linear mean-squared predictors of $\mathbf{X}_{t+h}$ and $\mathbf{Y}_{t+h}$ based on the observations up to time t are

$$\begin{aligned} P_t\mathbf{X}_{t+h} &= (F_{t+h-1}F_{t+h-2}\cdots F_{t+1})P_t\mathbf{X}_{t+1}, \\ P_t\mathbf{Y}_{t+h} &= G_{t+h}P_t\mathbf{X}_{t+h}. \end{aligned}$$

The reader is referred to Brockwell and Davis [2] for further information on structural modeling and prediction via the Kalman recursions.

We initially considered the relative differenced series to be a randomly varying trend with seasonal and noise components. This model was chosen because we wanted to determine whether there was a trend in the data and also whether there was a weekly effect. The state space form for this model is as above with $\mathbf{Y}_t$, $\mathbf{X}_t$, G_t, F_t, R_t and Q_t defined as follows:

$$\begin{aligned} \mathbf{Y}_t &= \frac{U_{t+1} - U_t}{U_t}, U_t = \text{maximum daily kw usage at time } t, \\ \mathbf{X}_t &= (M_t, B_t, s_t, s_{t-1}, s_{t-2}, s_{t-3}, s_{t-4}, s_{t-5})', \end{aligned}$$

where M_t is the deterministic signal at time t, B_t is the slope of the linear trend at time t and s_t is a seasonal component satisfying $s_{t+7} = s_t$ and $\sum_{t=1}^{7} s_t = 0$,

$$G_t = G = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$F_t = F = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$R_t = R = \sigma_w^2 I_{7\cdot 7}$$

$$Q_t = Q = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$t = 1, 2, \ldots, 183$ (time $t = 1$ representing April 1 and time $t = 183$ representing September 30).

This model was fit using a Fortran program which incorporates the use of Hooke and Jeeve's algorithm to find maximum likelihood estimators of σ_w^2 and Q via the Kalman recursions. $\hat{\mathbf{X}}_1 = 0$ and $\Omega_1 = 0 \cdot I_{7\cdot 7}$ were the initial conditions used in the fit. For further reference on obtaining these maximum likelihood estimates, see Brockwell and Davis [2].

From this initial fit, the series did not appear to have either a linear trend or a weekly component (both σ_2^2 and σ_3^2 were approximately 0). Therefore, we refit the model as a random walk with a noise term. This model fit the data quite well considering the fact that we had not yet incorporated any of the weather data. The observed values (denoted by 'o') and one-step predicted values (denoted by '*') of maximum kw usage under this model for 1992 through 1996 are depicted in Figures 21-25. The maximum likelihood estimates of σ_1^2 and σ_w^2 obtained by the Fortran program are given in Table 3.1.

Table 3.1: Maximum Likelihood Estimates

Year	σ_1^2	σ_w^2
1992	0.05	0.07
1993	0.05	0.06
1994	0.05	0.03
1995	0.05	0.03
1996	0.05	0.06

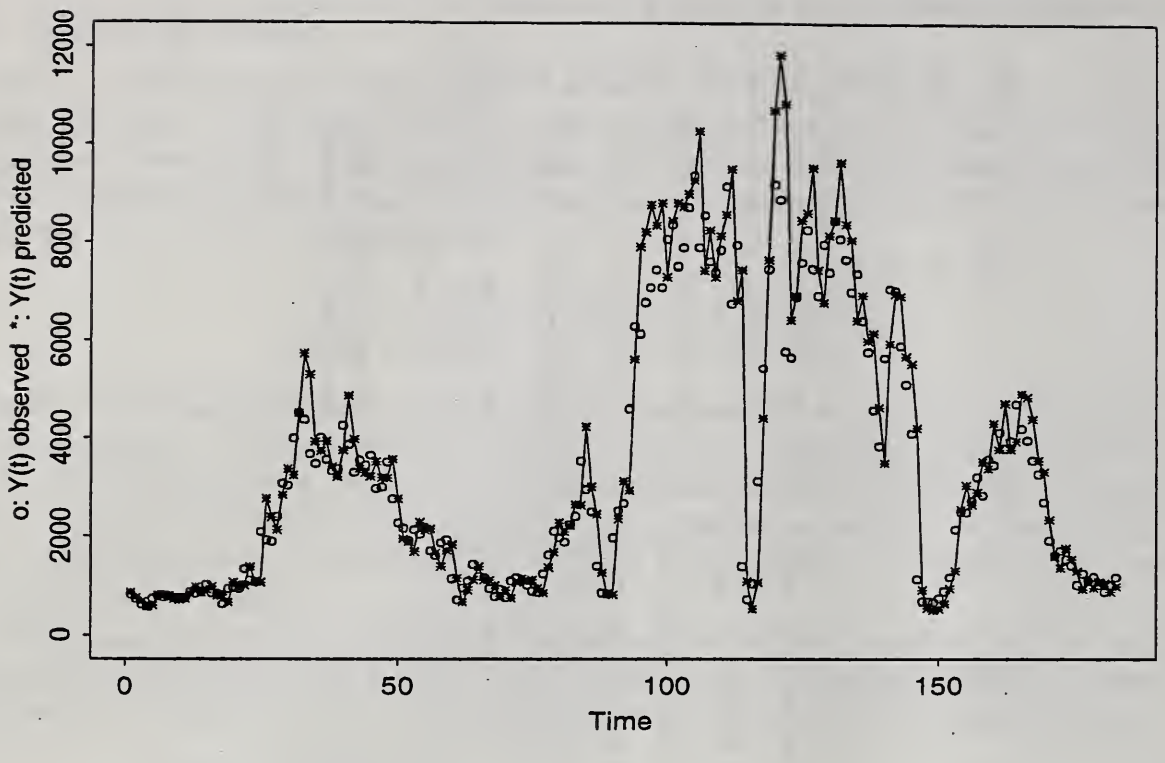

Figure 21: Structural Model Fit for 1992

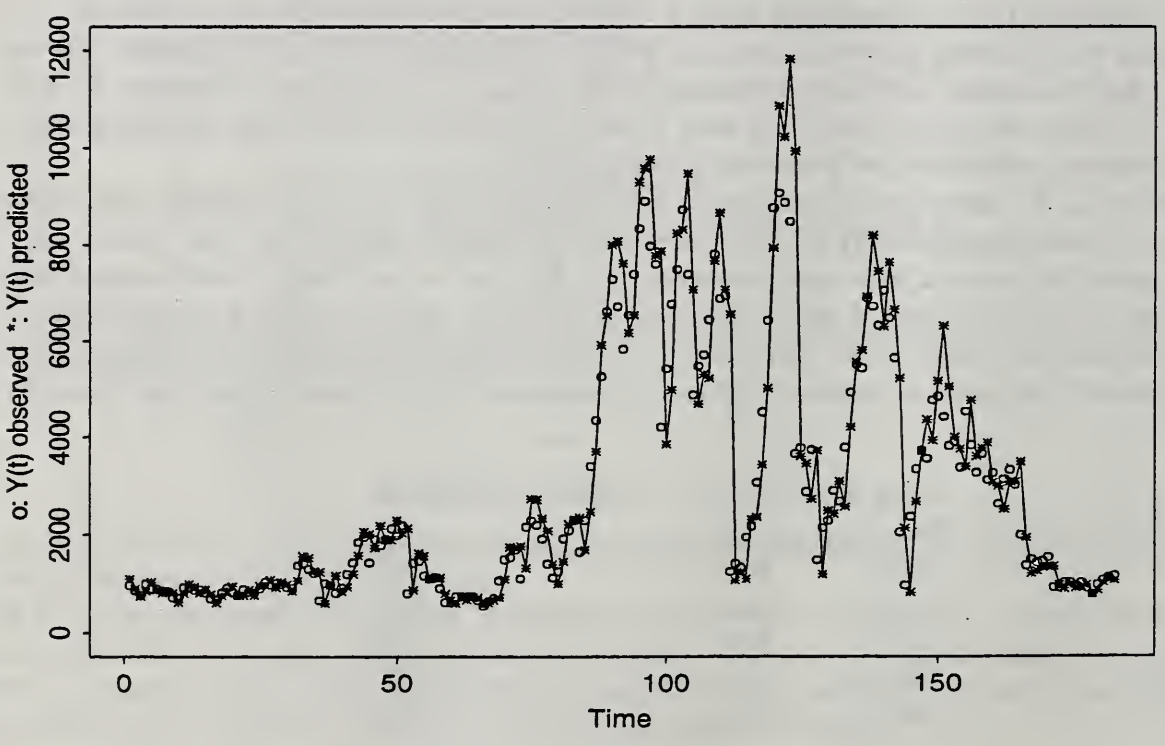

Figure 22: Structural Model Fit for 1993

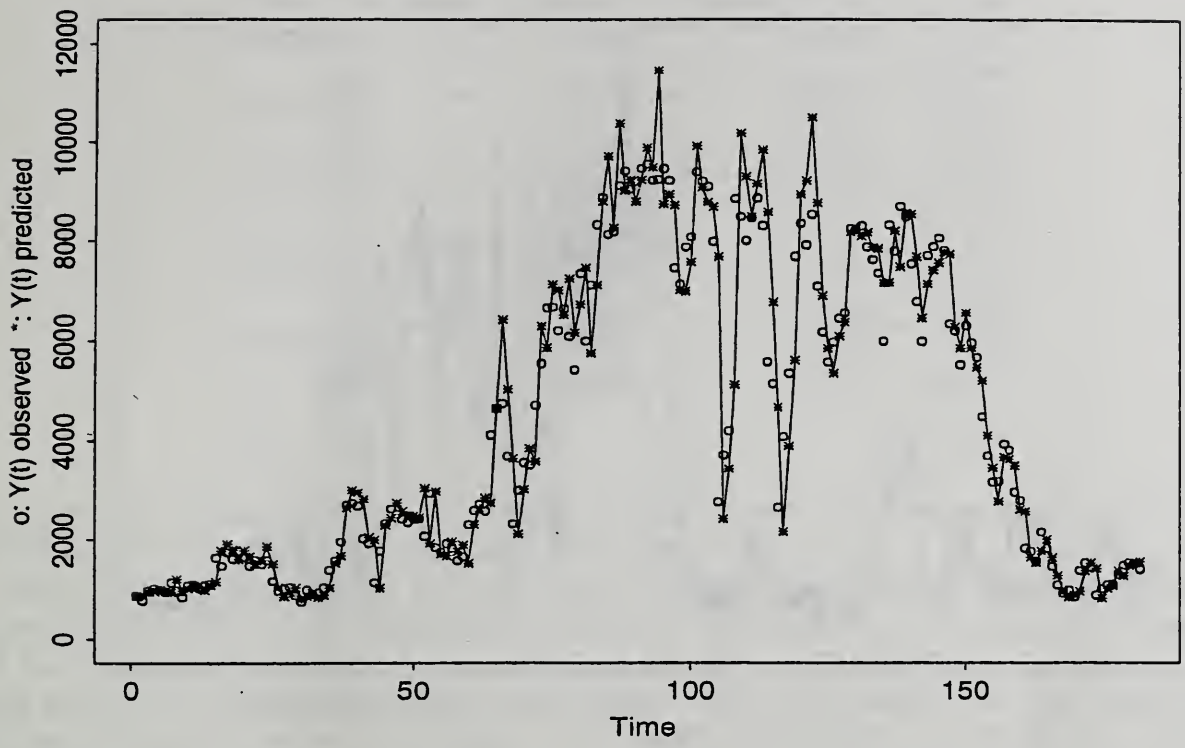

Figure 23: Structural Model Fit for 1994

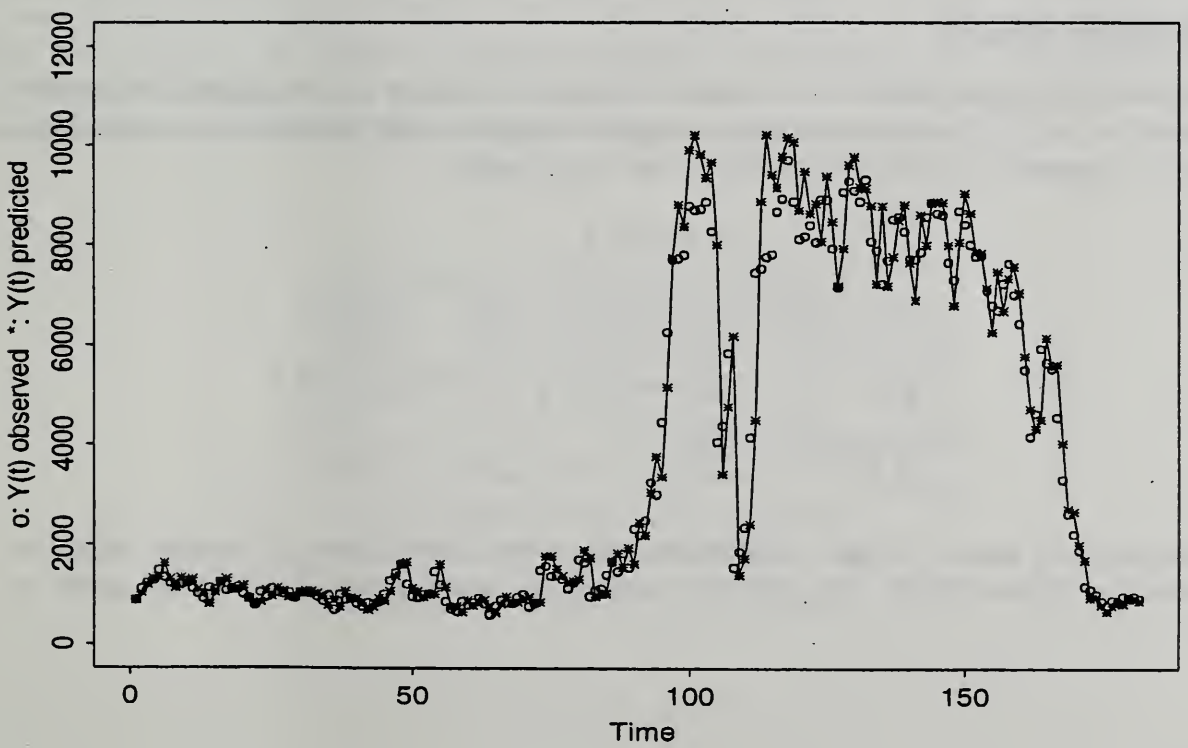

Figure 24: Structural Model Fit for 1995

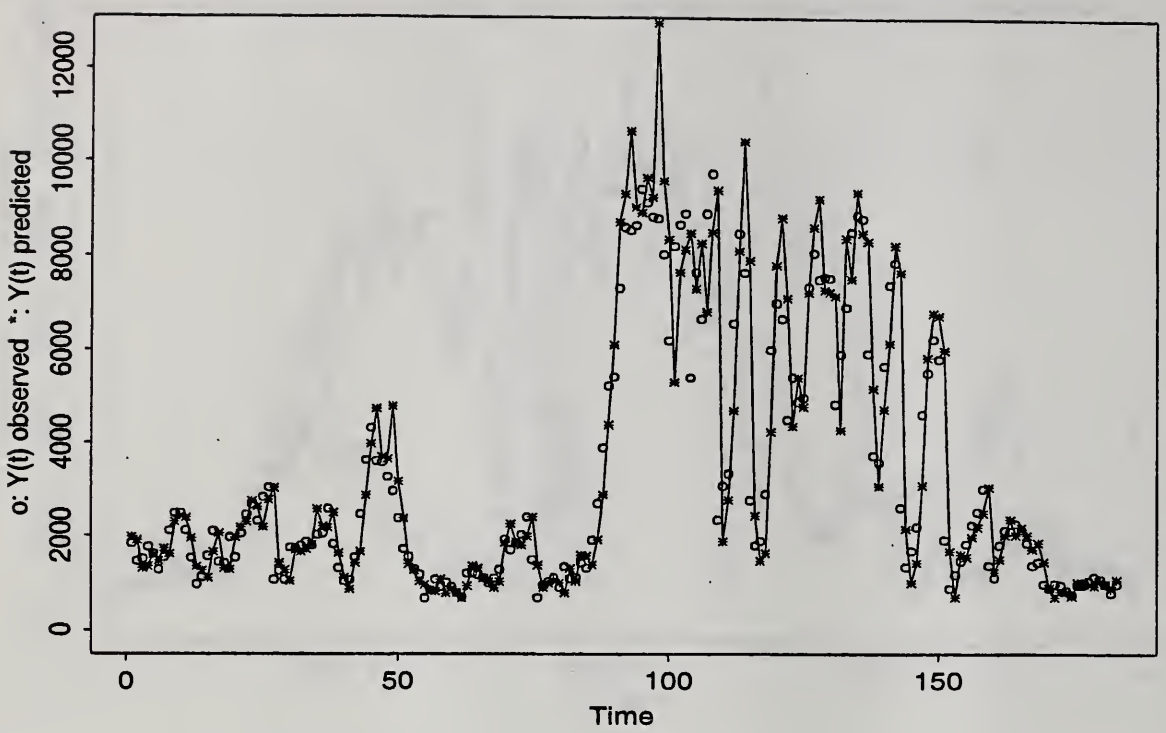

Figure 25: Structural Model Fit for 1996

3.2 Regression Models

In an effort to improve upon the structural model discussed in Section 3.1, we decided to consider various regression models to determine which weather variables might improve our predictions. One of the first regression models we considered was the following:

$$X_t = \beta_0 + \beta_1 R_t + \epsilon_t,$$

where:

$$\begin{aligned} X_t &= \frac{Y_{t+1} - Y_t}{Y_t}, \ Y_t = \text{maximum daily kw usage at time } t \\ R_t &= \text{daily rainfall at time } t \\ \epsilon_t &\sim N(0, \sigma^2) \end{aligned}$$

We considered this model because of the relationship between the relative differenced data and rainfall evident in Figures 16-20. The ANOVA table values obtained using S-Plus are given in Table 3.2a.

Table 3.2a: ANOVA Table Values

Year	Parameter	Estimate	Standard Error	P-value
1992	β_0	0.0702	0.0214	0.0013
	β_1	-0.0140	0.0029	0.0000
1993	β_0	0.0850	0.0193	0.0000
	β_1	-0.0217	0.0033	0.0000
1994	β_0	0.0410	0.0159	0.0109
	β_1	-0.0153	0.0041	0.0003
1995	β_0	0.0234	0.0150	0.1194
	β_1	-0.0058	0.0050	0.2422
1996	β_0	0.0735	0.0225	0.0013
	β_1	-0.0100	0.0024	0.0001

As we saw in Figures 16-20, the inverse relationship between the relative differenced data and rainfall is quite strong for all of the years except 1995.

Using this regression model we were able to obtain more accurate predictions (in terms of mean-squared error) for the original series Y_t than we were under the structural model framework. The root mean-squared errors of the predictions are given in Table 3.2b for both of these models as well as an additional regression model which will be discussed below. The mean-squared error for the original series was calculated using the formula $MSE = \frac{1}{182}\sum_{i=1}^{182}(Y_{t+1} - \hat{Y}_{t+1})^2 = \frac{1}{182}\sum_{i=1}^{182}(Y_{t+1} - Y_t(\hat{X}_t + 1))^2$. The sum consists of 182 terms instead of 183 terms since the models were fit to the relative differenced series $\{X_t\}$. The predictions for the series $\{Y_t\}$ were then found from the predictions of $\{X_t\}$ using the relationship $X_t = \frac{Y_{t+1} - Y_t}{Y_t}$. The gain in predictive power can also be seen by comparing Figures 21-25 with Figures 26-30. In these figures, the observed values are plotted as 'o' while the one-step predictions are plotted as '*'. The darkened circles in Figures 26-30 indicate that there was a positive amount of precipitation on the previous day. Again, the inverse relationship between power usage and rainfall is evident.

Table 3.2b: Root Mean-squared Errors

Model	1992	1993	1994	1995	1996
Model 1	983	1,016	870	745	1,216
Model 2	671	688	810	700	1,152
Model 3	717	753	819	720	1,195

Model 1 is the Structural Model, Model 2 is the Regression Model with Precipitation, and Model 3 is the Regression Model with Precipitation and ETA.

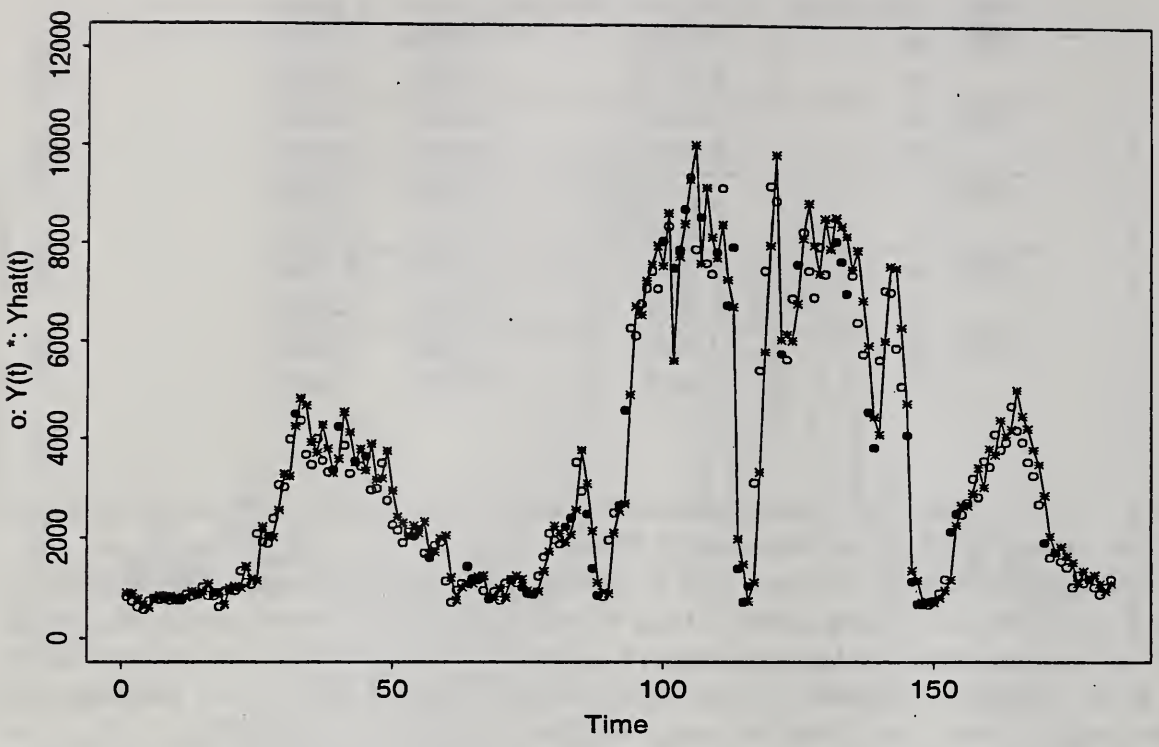

Figure 26: Regression Model with Precipitation - 1992

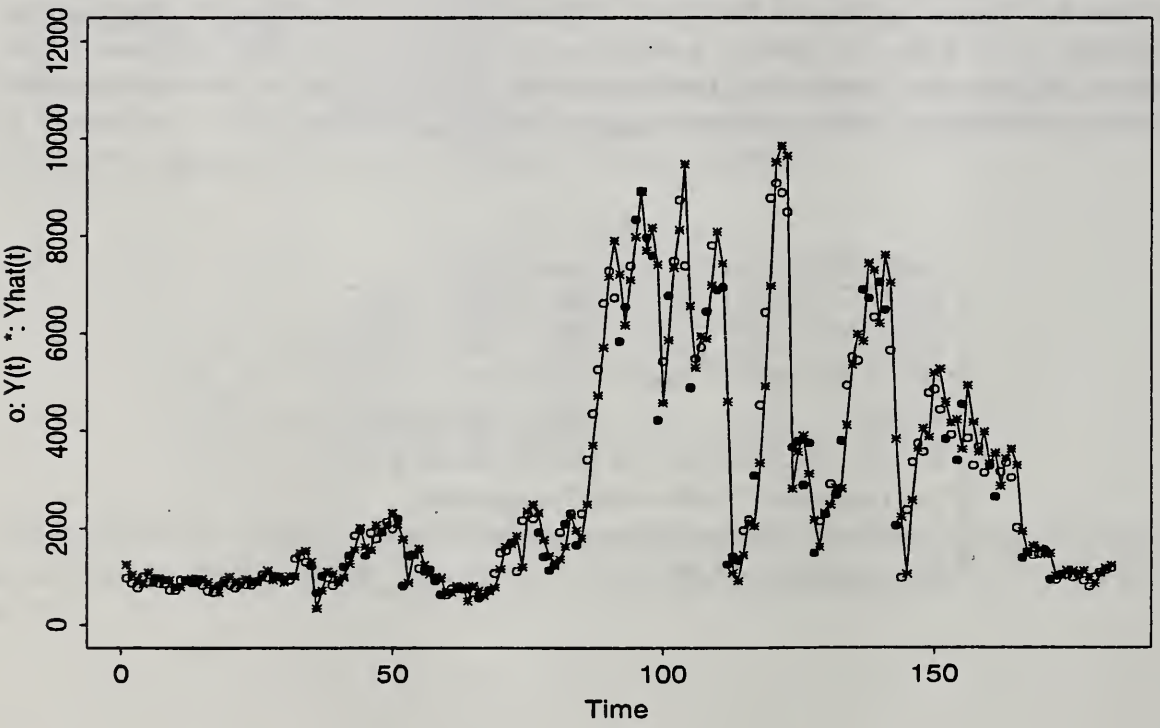

Figure 27: Regression Model with Precipitation - 1993

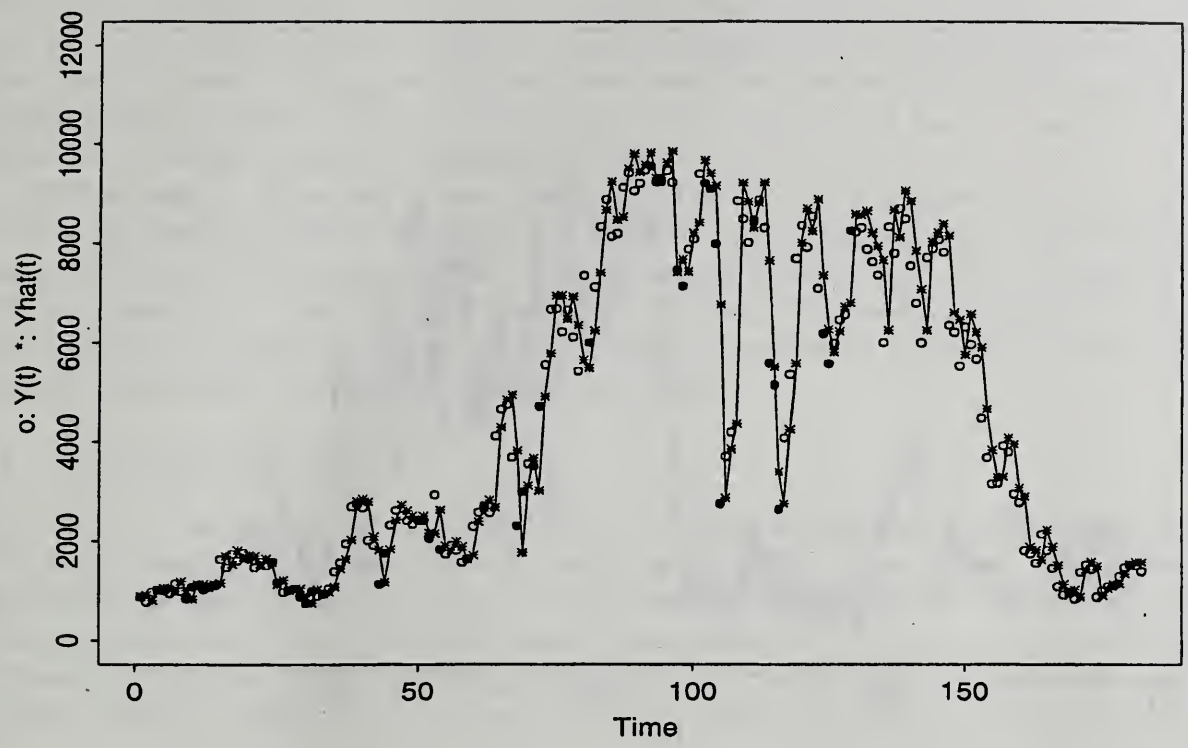

Figure 28: Regression Model with Precipitation - 1994

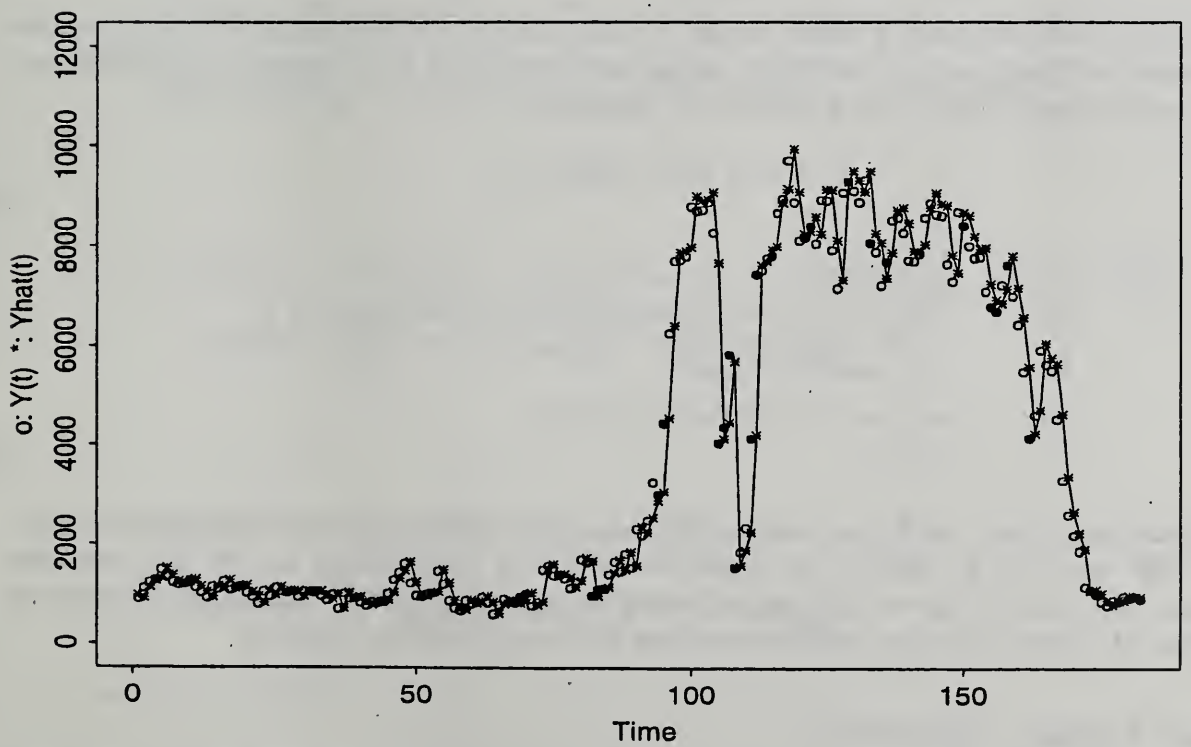

Figure 29: Regression Model with Precipitation - 1995

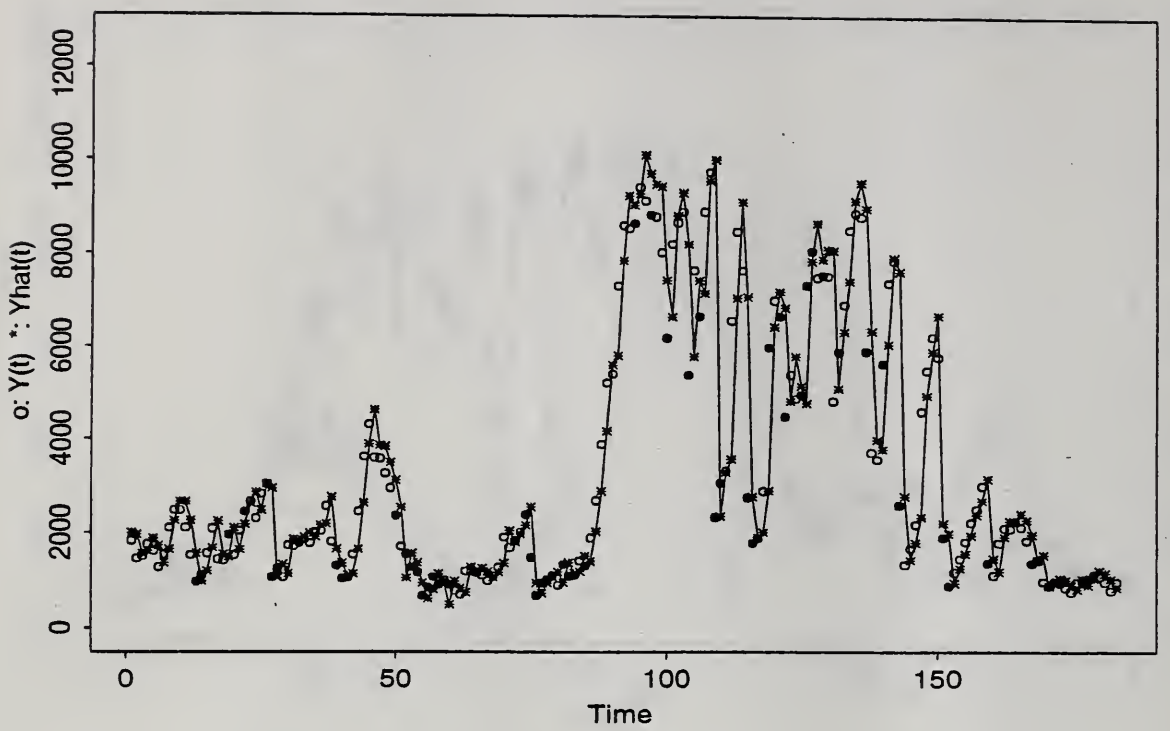

Figure 30: Regression Model with Precipitation - 1996

To determine whether other weather variables might assist in predicting power usage, we included various combinations of the available variables as predictors in a regression model setting. One such model is given below and is included in Table 3.2b.

$$X_t = \beta_0 + \beta_1 R_t + \beta_2 E_t + \epsilon_t$$

where:

$$\begin{aligned} X_t &= \frac{Y_{t+1} - Y_t}{Y_t}, \ Y_t = \text{maximum daily kw usage at time t} \\ R_t &= \text{daily rainfall at time t} \\ E_t &= \text{daily value of ETAcorn at time t} \\ \epsilon_t &\sim N(0, \sigma^2) \end{aligned}$$

The predictions under this model were better than those obtained from the structural model. However, they were not as accurate as those obtained with precipitation as the sole predictor variable (see Table 3.2b). This was the case with all of the regression models we considered involving combinations of rainfall and other weather variables as the explanatory variables.

4 Base Model Approach

At this point, we began to re-think our modeling strategy. In order to make predictions more than one day in advance with the regression model, we would be required to model rainfall, a difficult task. Also, as Figures 26-30 indicate, rainfall has different affects on power usage at different times

of the year. However, we are not able to easily incorporate this observation into a regression model framework. For these reasons, we decided to develop a base model which could be used as a starting point for a new year and whose coefficients could be updated as more information became available. Once the base model had been decided upon, we would then need to model its parameters. Having accomplished this task, we would be able to make predictions for the days remaining based on the information available at any previous point in time.

Our first step in developing a base model was to form a theoretical model which was suitable for all of the available years (1992-1996). We returned to modeling the daily maximum series $\{Y_t\}$ instead of the relative differenced series $\{X_t\}$, because the peaks of $\{X_t\}$ did not always correspond with the peaks of $\{Y_t\}$ and our main concern was to accurately model the peaks of the series of daily maxima. We began our search for a base model by examining sinusoidal models because the series $\{Y_t\}$ resembled a combination of sines and cosines.

4.1 Sinusoidal Models

Only the first three available years (1992-1994) were considered with somewhat discouraging results. We fit the sinusoidal models using absolute deviation regression (L_1 regression). This procedure fits a regression model by minimizing the sum of the absolute deviations of the errors instead of the sum of the squared errors. This type of loss function was chosen in an effort to obtain predictions that were more robust than those found using a squared error loss function (L_2 regression). Figures 31-33 depict both the observed ('o') and predicted ('*') values for the fitted models given below.

1992:

$$\begin{aligned} Y_t &= 3724.98 - 788.11\cos(\frac{\pi t}{15}) - 1288.28\cos(\frac{\pi t}{30}) - 1641.94\cos(\frac{\pi t}{60}) + 428.04\sin(\frac{\pi t}{10}) \\ &\quad +383.80\sin(\frac{\pi t}{15}) - 2024.10\sin(\frac{\pi t}{35}) + 1999.57\sin(\frac{\pi t}{45}) - 3422.79\sin(\frac{\pi t}{84}) \end{aligned}$$

1993:

$$\begin{aligned} Y_t &= 2931.96 - 57.70\cos(\frac{\pi t}{8}) - 1201.66\cos(\frac{\pi t}{30}) - 949.12\cos(\frac{\pi t}{85}) + 88.89\sin(\frac{\pi t}{15}) \\ &\quad +790.69\sin(\frac{\pi t}{21}) - 681.95\sin(\frac{\pi t}{30}) - 2203.08\sin(\frac{\pi t}{84}) \end{aligned}$$

1994:

$$\begin{aligned} Y_t &= 4140.73 - 151.29\cos(\frac{\pi t}{19}) - 878.33\cos(\frac{\pi t}{30}) + 593.62\cos(\frac{\pi t}{50}) - 2733.19\cos(\frac{\pi t}{105}) \\ &\quad +178.98\sin(\frac{\pi t}{6}) - 2209.81\sin(\frac{\pi t}{84}) + 708.34\sin(\frac{\pi t}{105}) \end{aligned}$$

$$t = 1, 2, \ldots, 183.$$

Only two of the predictor variables were the same for all years, and minor changes in the explanatory variables greatly varied the fit. This made it difficult to arrive at a base model that would be suitable for all of the years. Therefore, this type of model received no further consideration.

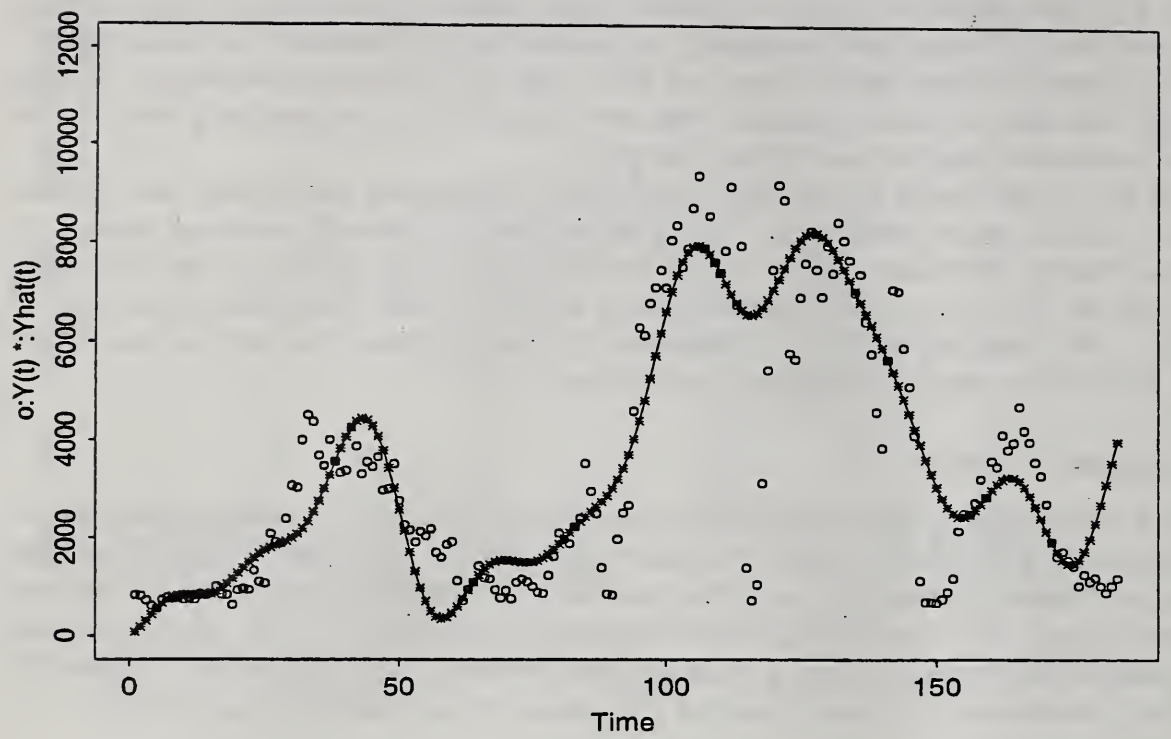

Figure 31: Sinusoidal Model - 1992

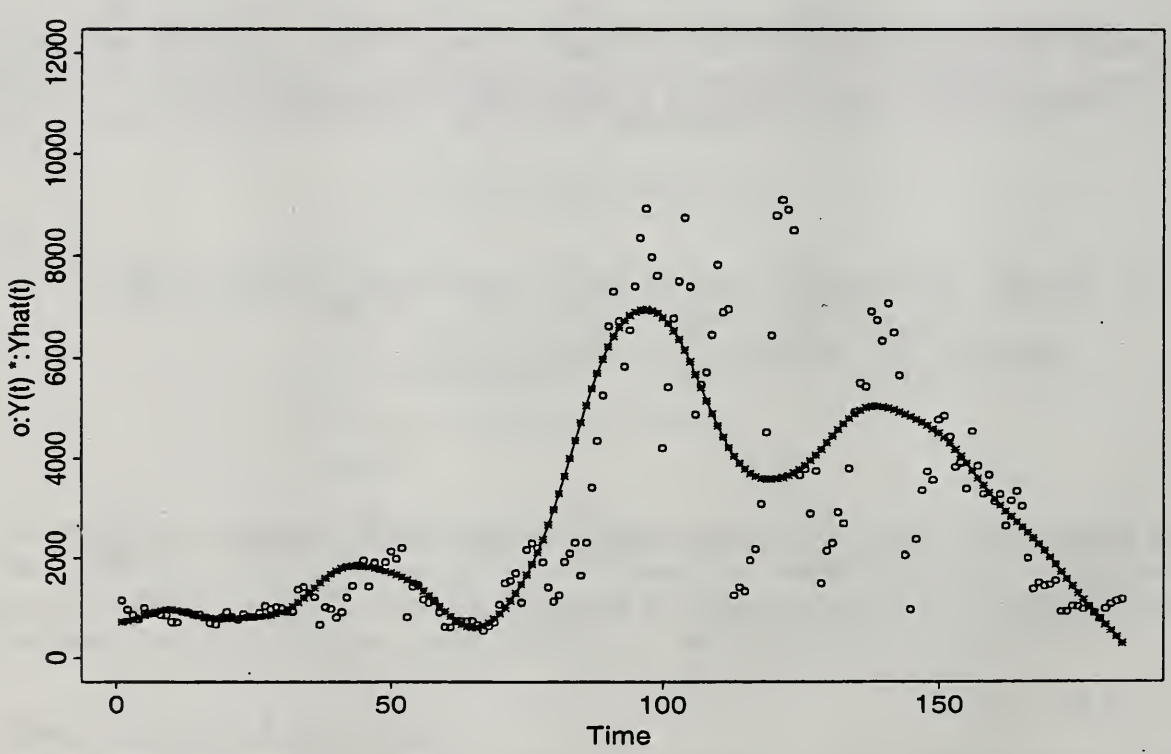

Figure 32: Sinusoidal Model - 1993

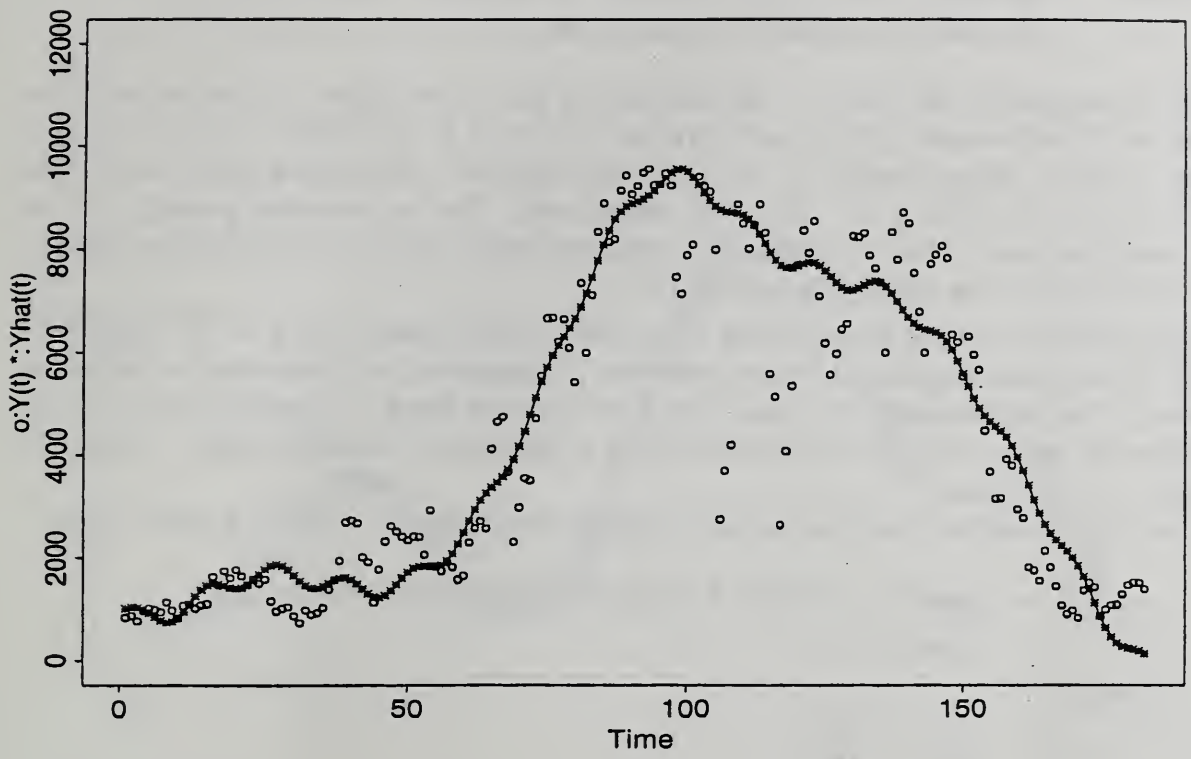

Figure 33: Sinusoidal Model - 1994

4.2 Polynomial Models

We next considered the use of polynomial functions as our predictors in the base model. The polynomials we decided upon were shifted Legendre polynomials.

4.2.1 Shifted Legendre Polynomials

One of the reasons these types of polynomials were considered is because they are orthogonal on the interval from 0 to 1. The first ten shifted Legendre polynomials are given below.

$$
\begin{aligned}
f_0(x) &= 1 \\
f_1(x) &= 2x - 1 \\
f_2(x) &= 6x^2 - 6x + 1 \\
f_3(x) &= 20x^3 - 30x^2 + 12x - 1 \\
f_4(x) &= 70x^4 - 140x^3 + 90x^2 - 20x + 1 \\
f_5(x) &= 252x^5 - 630x^4 + 560x^3 - 210x^2 + 30x - 1 \\
f_6(x) &= 924x^6 - 2,772x^5 + 3,150x^4 - 1,680x^3 + 420x^2 - 42x + 1 \\
f_7(x) &= 3,432x^7 - 12,012x^6 + 16,632x^5 - 11,550x^4 + 4,200x^3 - 756x^2 + \\
&\quad 56x - 1 \\
f_8(x) &= 12,870x^8 - 51,480x^7 + 84,084x^6 - 72,072x^5 + 34,650x^4 - \\
&\quad 9,240x^3 + 1,260x^2 - 72x + 1
\end{aligned}
$$

$$f_9(x) = 48,620x^9 - 218,790x^8 + 411,840x^7 - 420,420x^6 + 252,252x^5 - 90,090x^4 + 18,480x^3 - 1,980x^2 + 90x - 1$$

Although we originally used these ten polynomials, we later scaled them in order to obtain an orthonormal set of polynomials $\{p_i\}$ ($\int_0^1 p_i(x)p_j(x)dx = 0$ for $i \neq j$; $\int_0^1 p_i^2(x)dx = 1$). The scaling factors used to ensure orthonormality for the ten shifted Legendre polynomials given above were 1, $\sqrt{3}$, $\sqrt{5}$, $\sqrt{7}$, 3, $\sqrt{11}$, $\sqrt{13}$, $\sqrt{15}$, $\sqrt{17}$, $\sqrt{19}$, respectively. The orthonormal property is nice because it simplifies many of the calculations involving the design matrix. The estimates given in this section are based on the scaled polynomials $\{p_i\}$.

Models consisting of various combinations of the scaled polynomials with $x = \frac{\text{day of year-90}}{183}$ were initially fit using both absolute deviation regression (L_1 regression) and least-squares regression (L_2 regression). The predictions of the peaks were more accurate using L_1 regression for all of the years. This was to be expected since the severe drops in usage have less of an impact under this loss function.

The model which was the most parsimonious without losing predictive power is given below.

$$Y(t) = \beta_0 p_0(t) + \beta_1 p_1(t) + \beta_2 p_2(t) + \beta_3 p_3(t) + \beta_4 p_5(t) + \beta_5 p_7(t) + \beta_6 p_9(t) + \epsilon(t)$$

where

$$t = \frac{1}{183}, \ldots, \frac{183}{183}.$$

Table 4.2.1 contains the parameter estimates and standard deviations under this model for the L_1 regression fit for the years 1992-1996. Figures 34-38 depict both the observed ('o') and predicted ('*') values for these years using this model.

The use of shifted Legendre polynomials resulted in a more parsimonious model and easier identification of each polynomial's contribution to the overall fit than did some of the other types of orthogonal polynomials which were considered. The advantage of identifying each polynomial's contribution will be discussed in more detail in Section 5.1.

Table 4.2.1: Parameter Estimates Using Shifted Legendre Polynomials: L_1 Regression

Parameter	1992	1993	1994	1995	1996
β_0	3437.28	2820.11	4267.23	3946.93	3145.48
	(51.34)	(48.41)	(61.12)	(70.38)	(66.98)
β_1	861.62	1045.87	1219.31	1834.97	679.90
	(51.31)	(48.39)	(61.09)	(70.34)	(66.94)
β_2	-1295.31	-1067.69	-2295.49	-1089.79	-1098.90
	(51.50)	(48.56)	(61.31)	(70.60)	(67.19)
β_3	-940.87	-1123.96	-1175.28	-2097.92	-1290.70
	(51.31)	(48.38)	(61.09)	(70.34)	(66.94)
β_4	957.84	617.87	576.30	529.66	890.92
	(51.26)	(48.34)	(61.03)	(70.28)	(66.88)
β_5	-1033.36	-350.58	377.70	-300.39	-362.90
	(51.16)	(48.24)	(60.90)	(70.13)	(66.74)
β_6	563.55	344.79	-29.59	464.82	224.52
	(50.96)	(48.05)	(60.67)	(69.86)	(66.48)

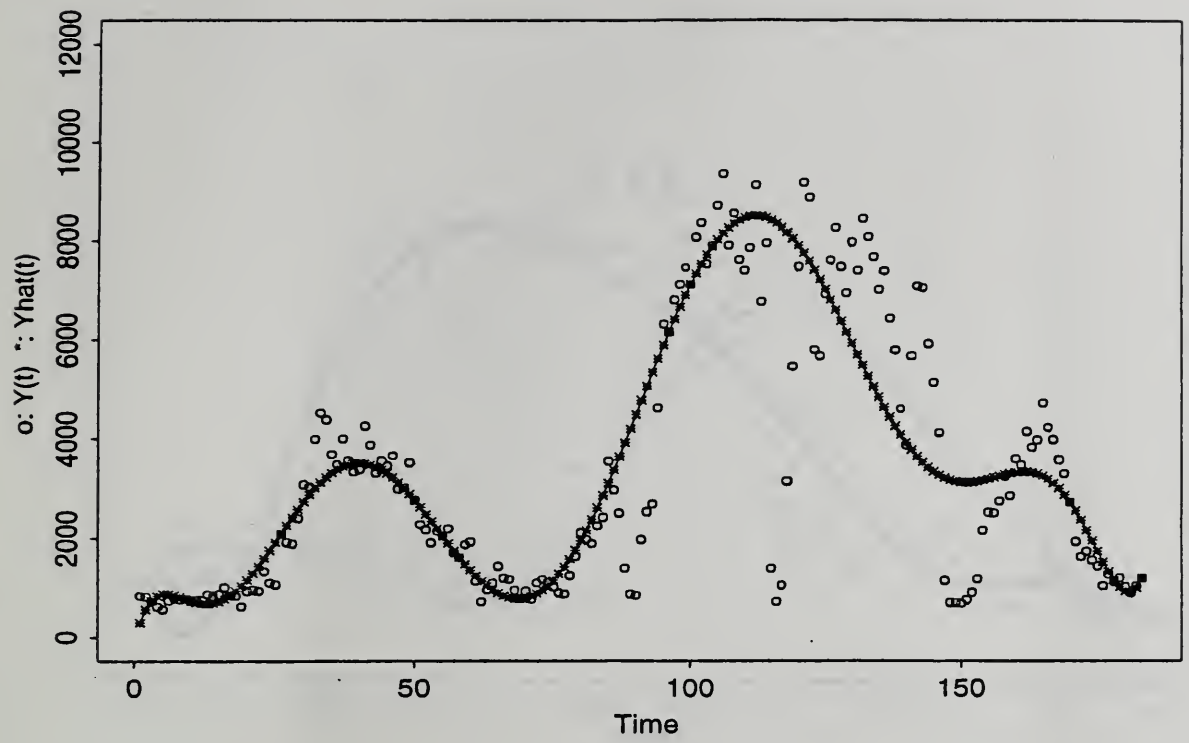

Figure 34: L_1 Legendre Polynomial Regression - 1992

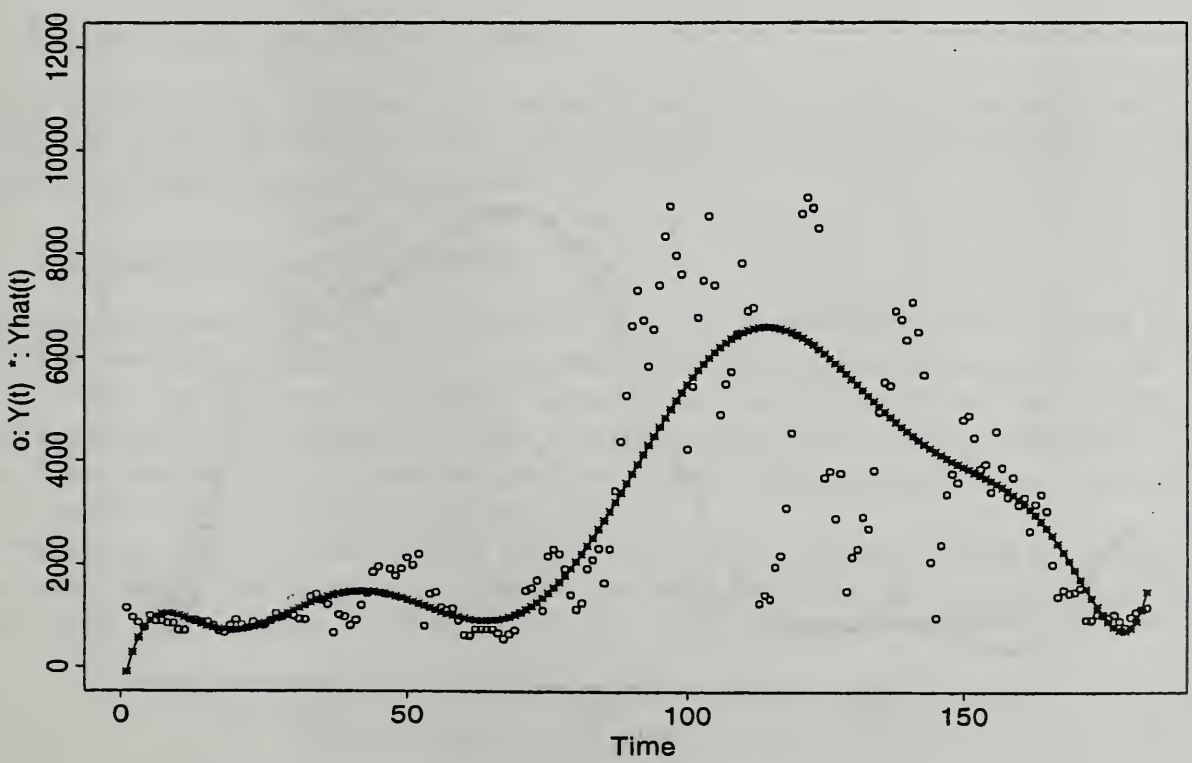

Figure 35: L_1 Legendre Polynomial Regression - 1993

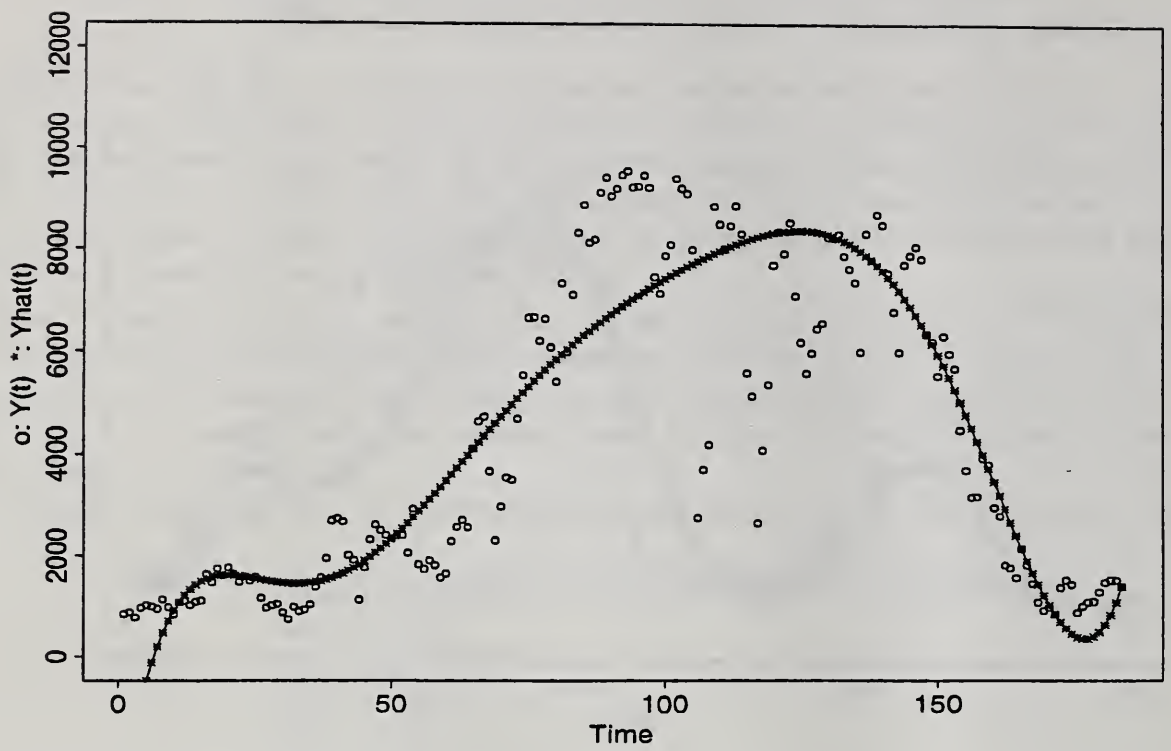

Figure 36: L_1 Legendre Polynomial Regression - 1994

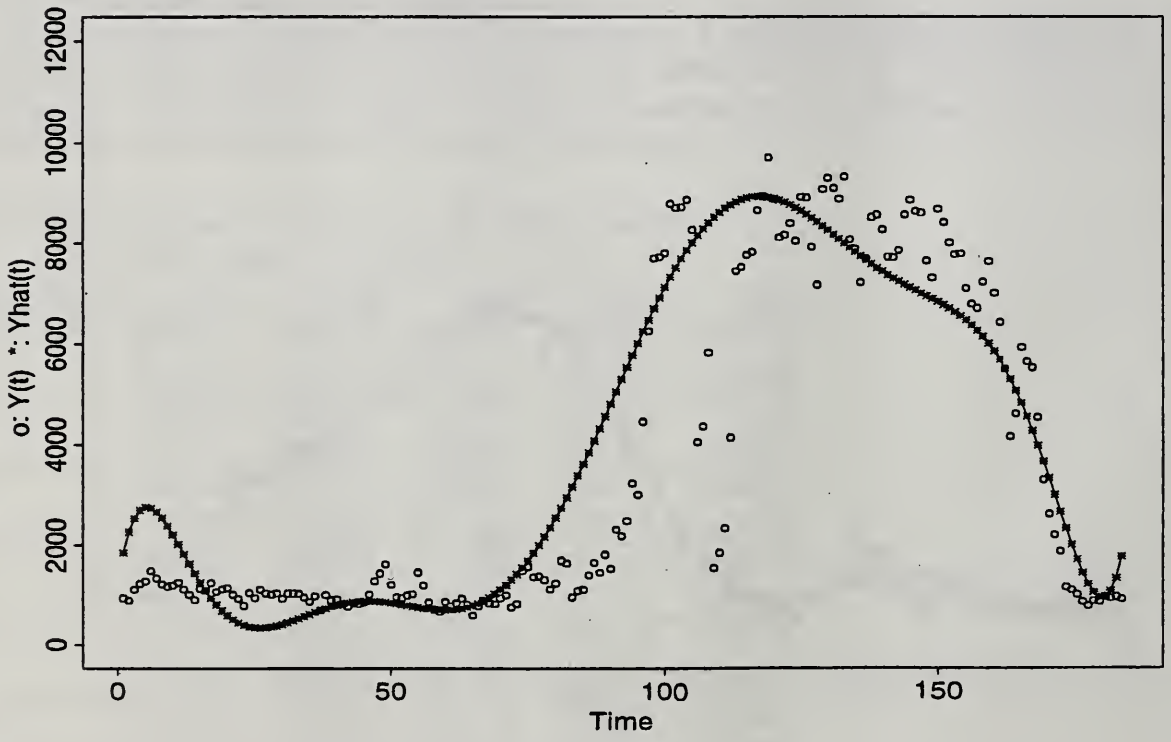

Figure 37: L_1 Legendre Polynomial Regression - 1995

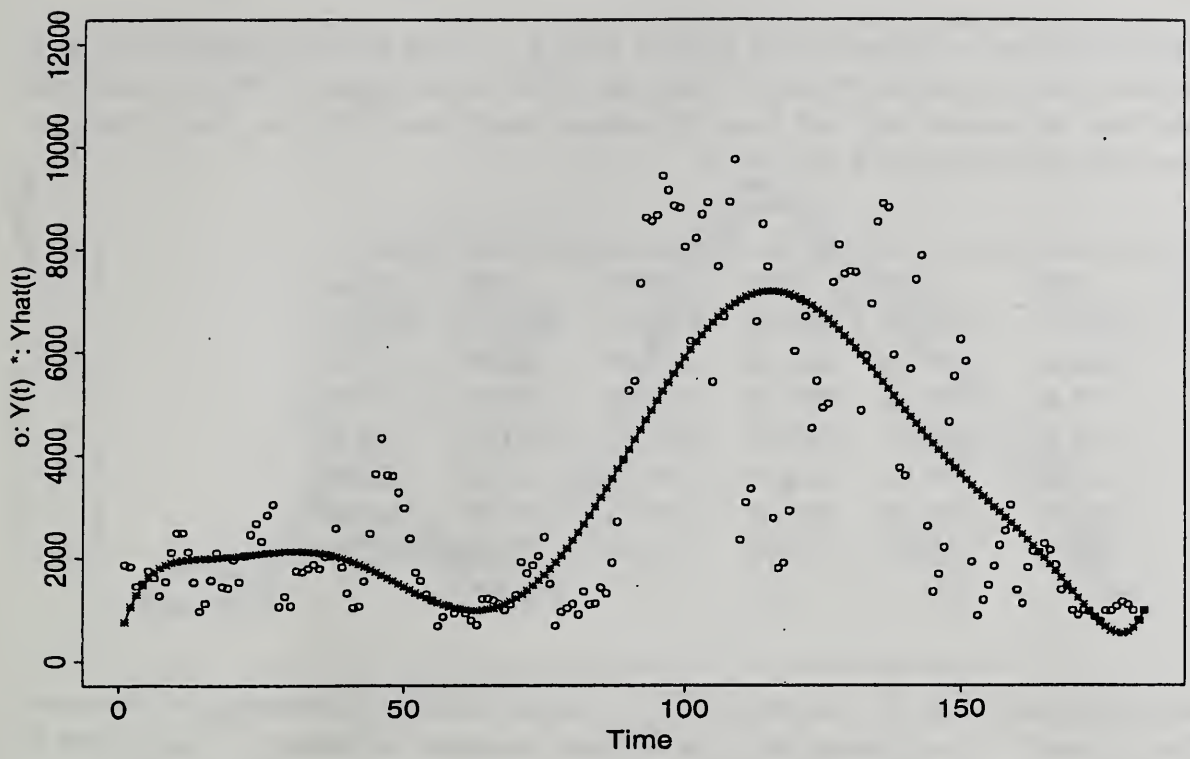

Figure 38: L_1 Legendre Polynomial Regression - 1996

5 Improving the Base Model

Having decided which orthogonal polynomials to use in the base model, we began to refine the model. We started by trying to obtain more accurate predictions of the peaks in the data by considering different types of loss functions.

5.1 Asymmetric Loss Functions

Although predictions of the peaks were better using L_1 regression than they were using L_2 regression, they were still not satisfactory. The problem with using this type of loss function was that all of the extreme points received less weight, including the peaks. Therefore, we began to explore the use of asymmetric loss functions, functions that give more weight to positive residuals. Use of this type of loss function results in more accurate predictions of the peaks (as well as poorer predictions of the valleys, but this was not of concern to us).

After many trials, the loss function we decided to use at this stage was an absolute deviation loss which gave double weight to the positive residuals. Using this loss function, we fit the same trend model as in Section 4.2.1. This model is reproduced below.

$$\begin{aligned} Y(t) &= \beta_0 p_0(t) + \beta_1 p_1(t) + \beta_2 p_2(t) + \beta_3 p_3(t) + \beta_4 p_5(t) + \beta_5 p_7(t) + \\ &\quad \beta_6 p_9(t) + \epsilon(t), \end{aligned}$$

where:

$$\begin{aligned} p_i(t) &= \text{the ith scaled, shifted Legendre polynomial at time t, and} \\ t &= \frac{1}{183}, \ldots, \frac{183}{183}. \end{aligned}$$

The parameter estimates for this model are given in Table 5.1. Along with the estimates for each year, the average of all five years is included in the table. With the exception of 1994, the estimates for each year were of the same sign and close to the same magnitude. This was the property we had been searching for in creating a base model.

Table 5.1: Parameter Estimates Based on an Asymmetric Loss Function

Parameter	1992	1993	1994	1995	1996	average
β_0	3855.73	3385.58	4682.06	4423.81	3833.92	4036.22
β_1	1069.31	1302.74	1317.00	1878.50	820.47	1277.61
β_2	-1378.47	-1385.32	-2204.87	-1020.76	-1449.68	-1487.82
β_3	-1125.87	-1383.07	-1216.69	-2335.39	-1561.61	-1524.53
β_4	1018.92	747.14	425.41	507.56	1145.80	768.97
β_5	-1079.67	-376.33	332.89	-427.71	-492.58	-408.68
β_6	509.18	358.21	-30.30	455.44	293.01	317.11

Plots of the observed values, 'o', and both sets of the predicted values (those using the parameter estimates for the year, '*' and those using the average parameter estimates, '^') are given in Figures 39 through 43. Use of this loss function resulted in more accurate predictions of the peaks than did the previously considered symmetric loss functions. Additionally, with the understanding that the estimates were not meant to be used as predictions, simply as our best guess for a new year, the average estimates were a promising start for our base model.

Plots of the polynomials p_i, i=1,2,3,5,7,9, times their estimated coefficients $\hat{\beta}_j$, j=1,...,6, are given in Figures 44 through 49. These graphs show the contribution of each polynomial to the overall fit of the model. With a few exceptions, each polynomial's contribution is similar from one year to the next. Again, this reinforced our belief that we had found an adequate base model.

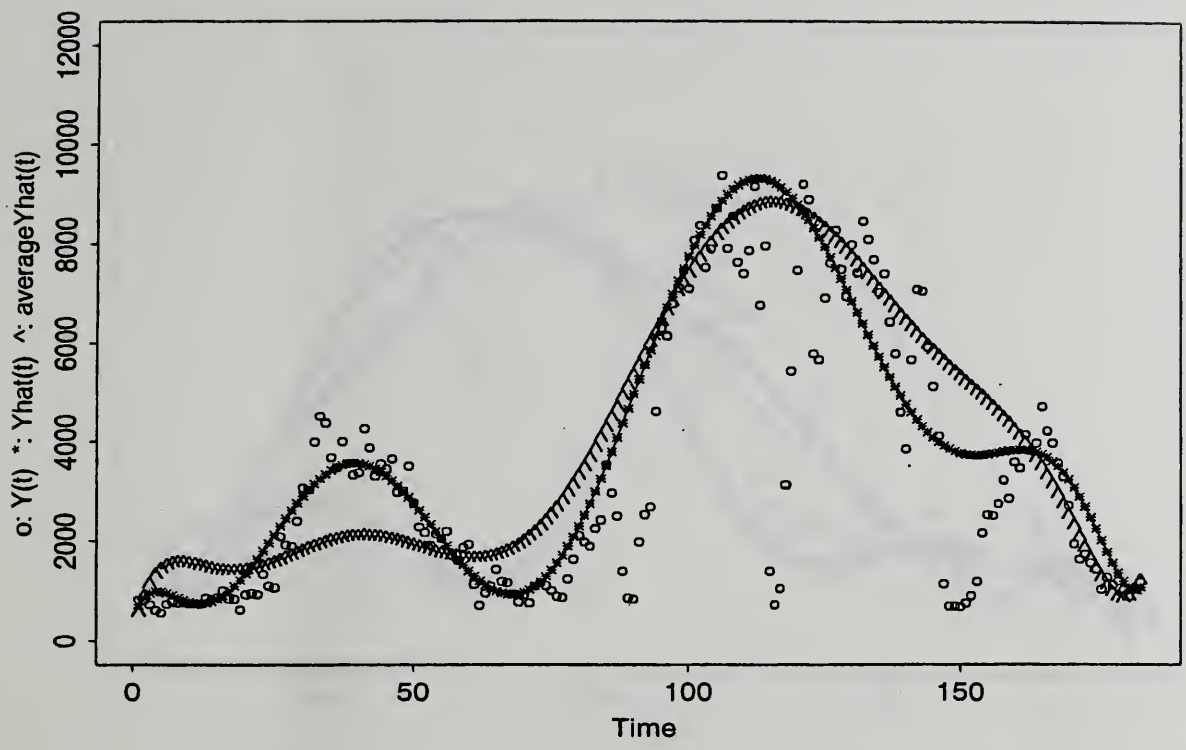

Figure 39: Asymmetric Loss Function- 1992

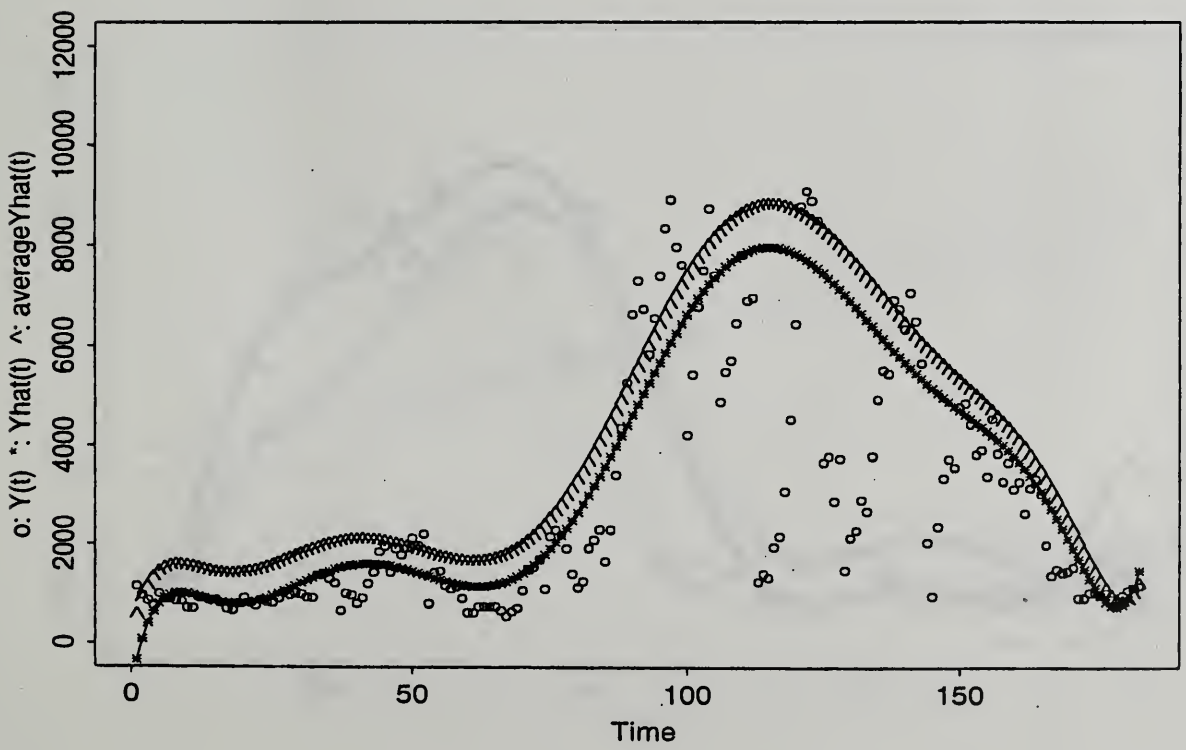

Figure 40: Asymmetric Loss Function- 1993

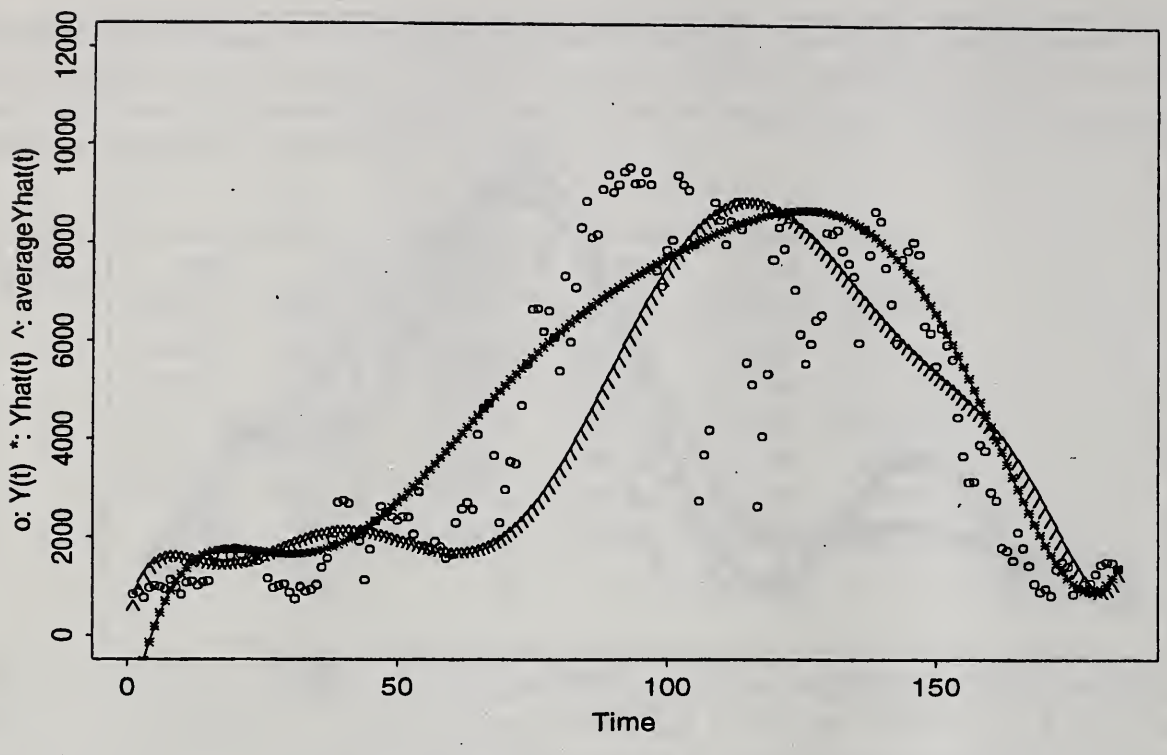

Figure 41: Asymmetric Loss Function- 1994

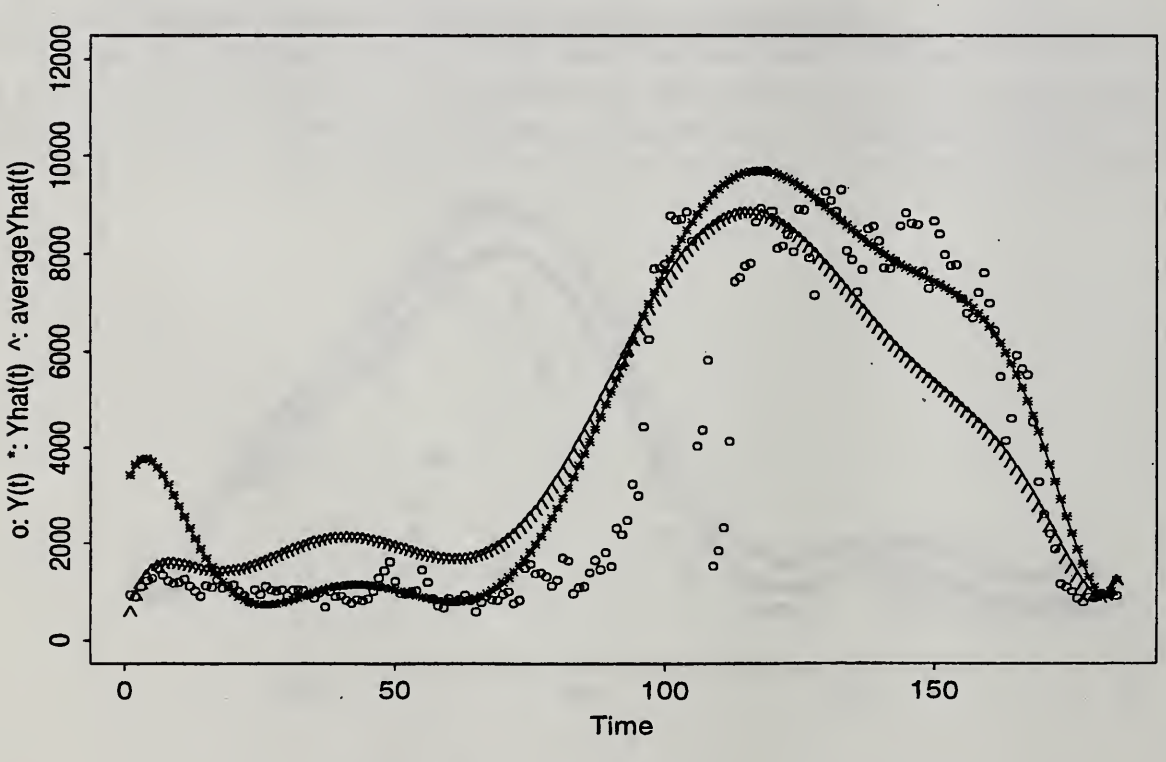

Figure 42: Asymmetric Loss Function- 1995

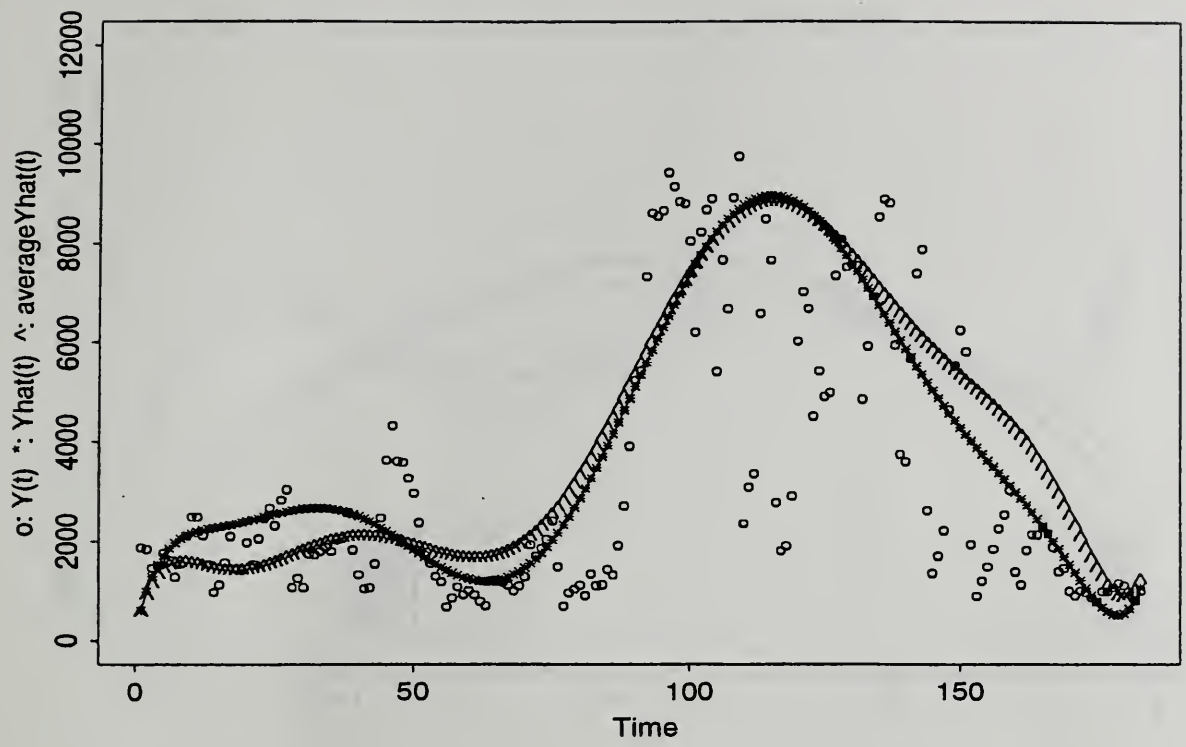

Figure 43: Asymmetric Loss Function- 1996

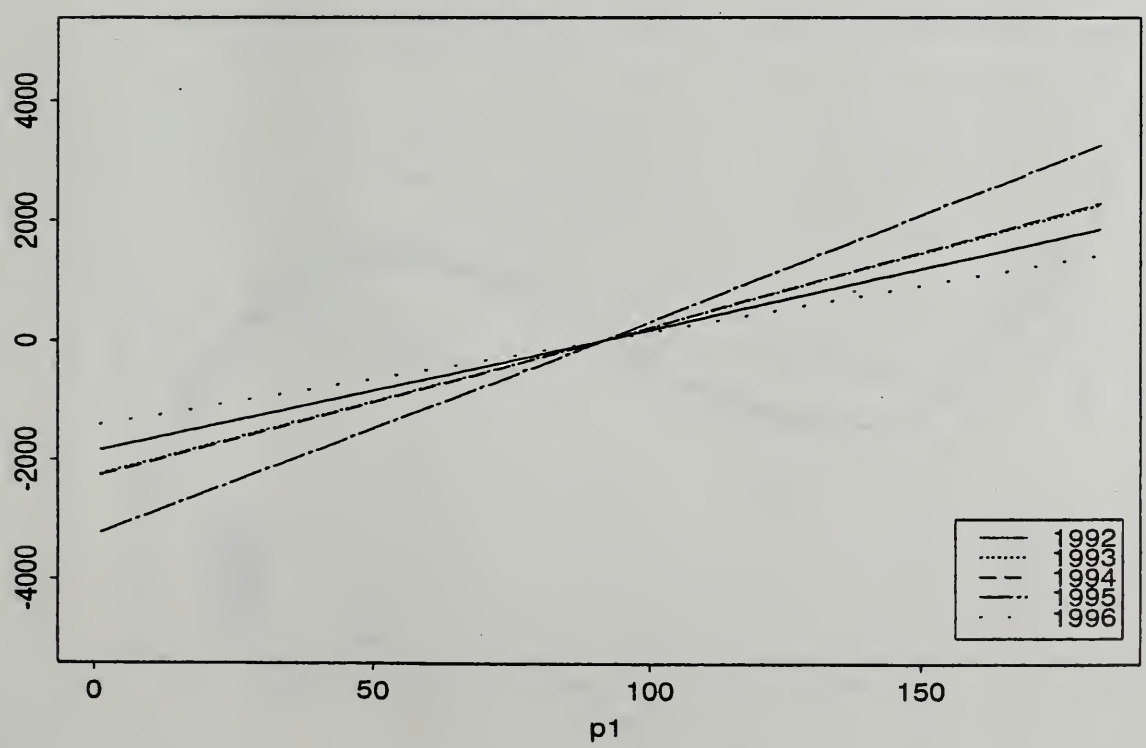

Figure 44: $\hat{\beta}_1 * p_1$

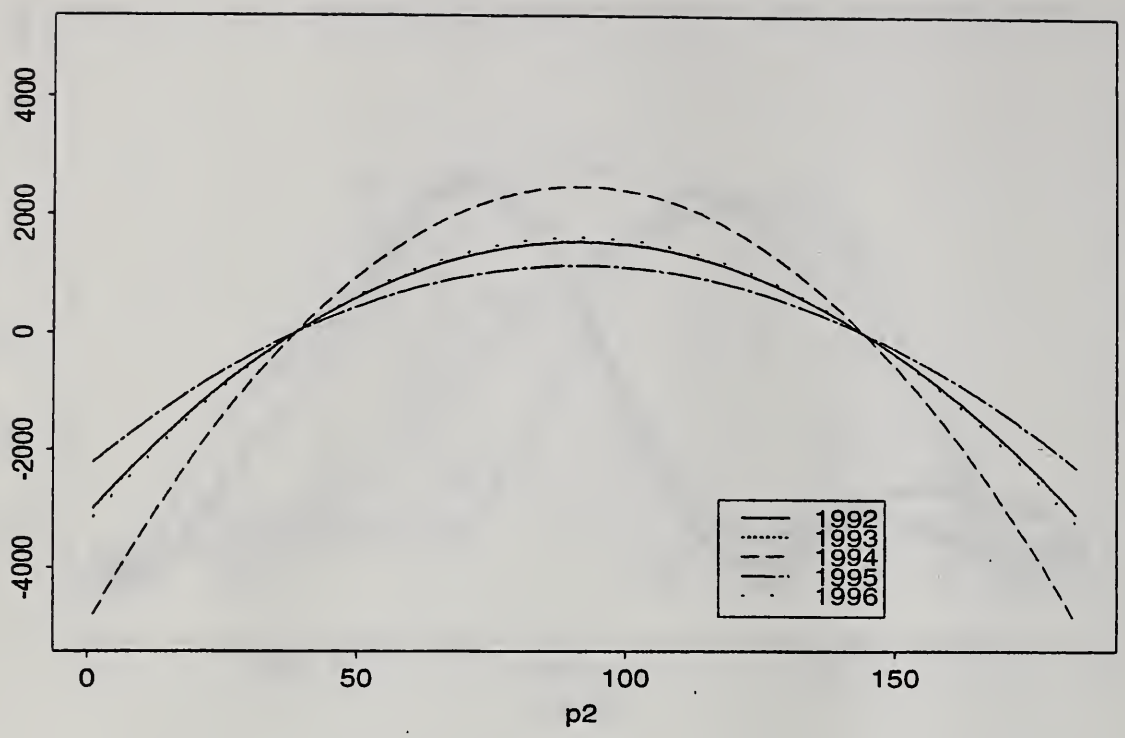

Figure 45: $\hat{\beta}_2 * p_2$

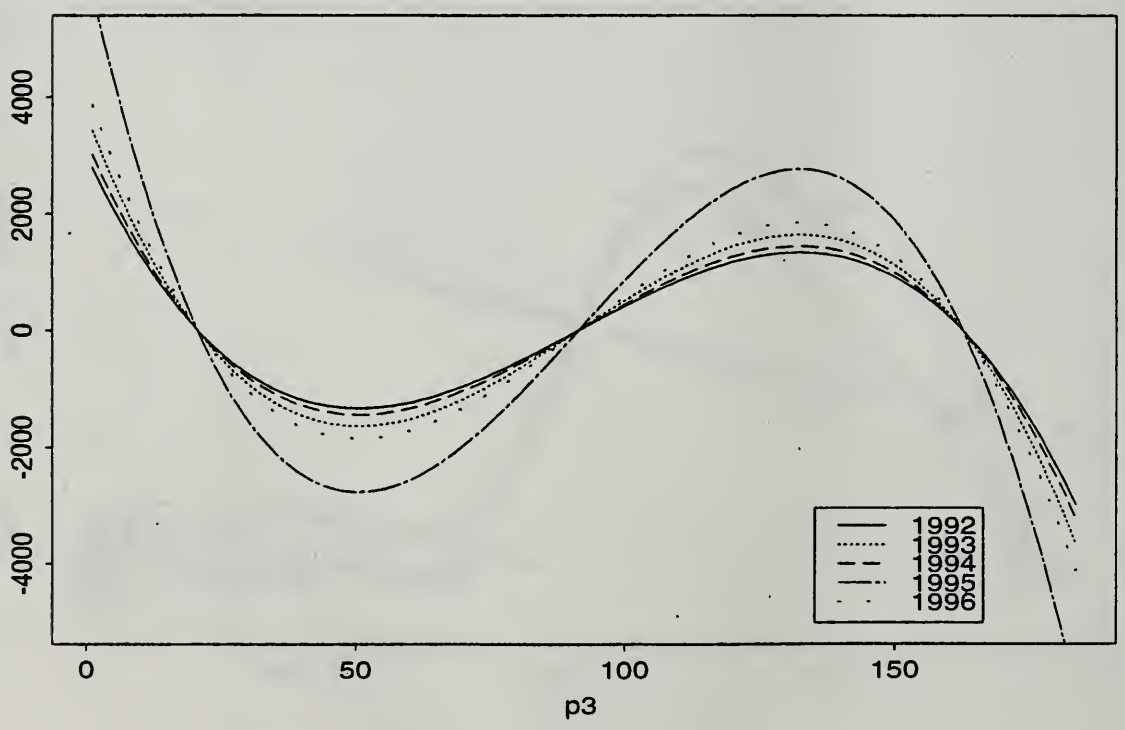

Figure 46: $\hat{\beta}_3 * p_3$

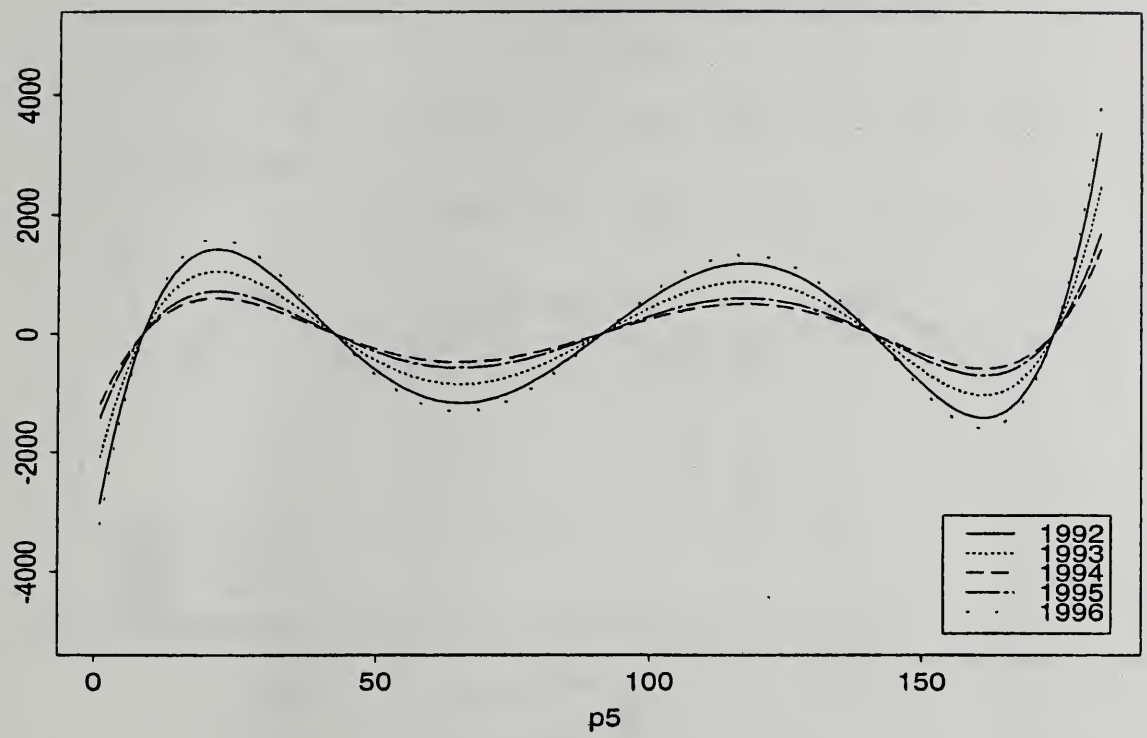

Figure 47: $\hat{\beta}_4 * p_5$

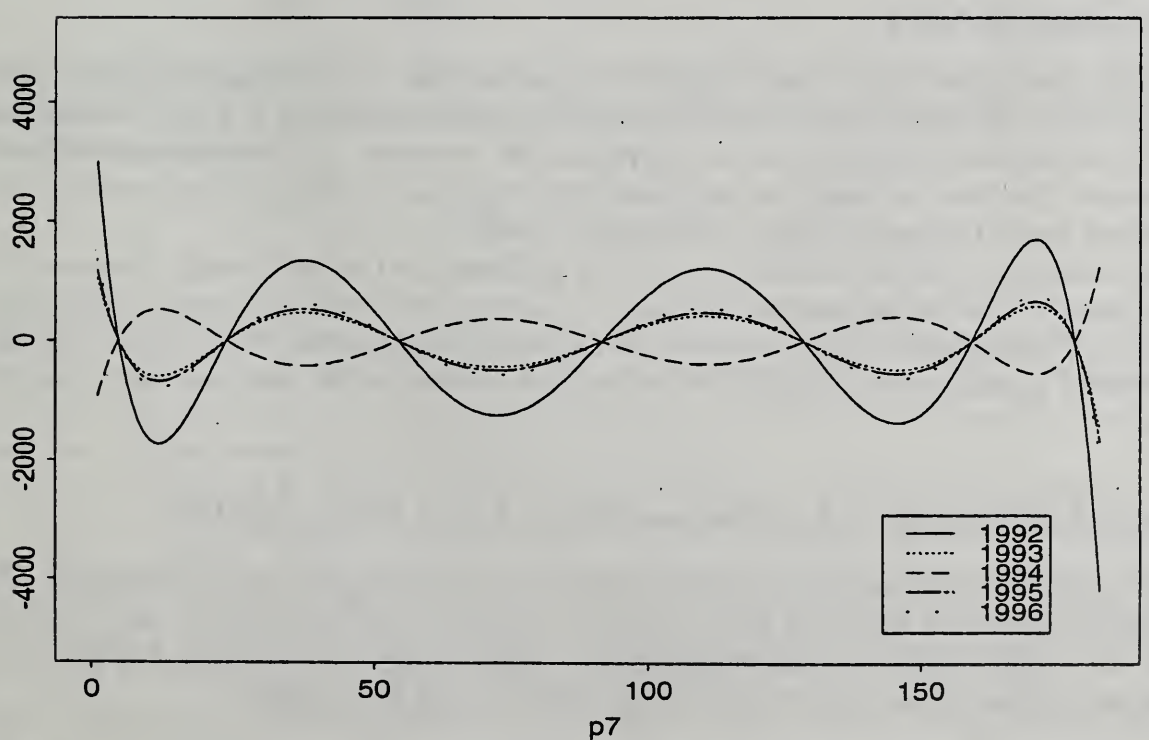

Figure 48: $\hat{\beta}_5 * p_7$

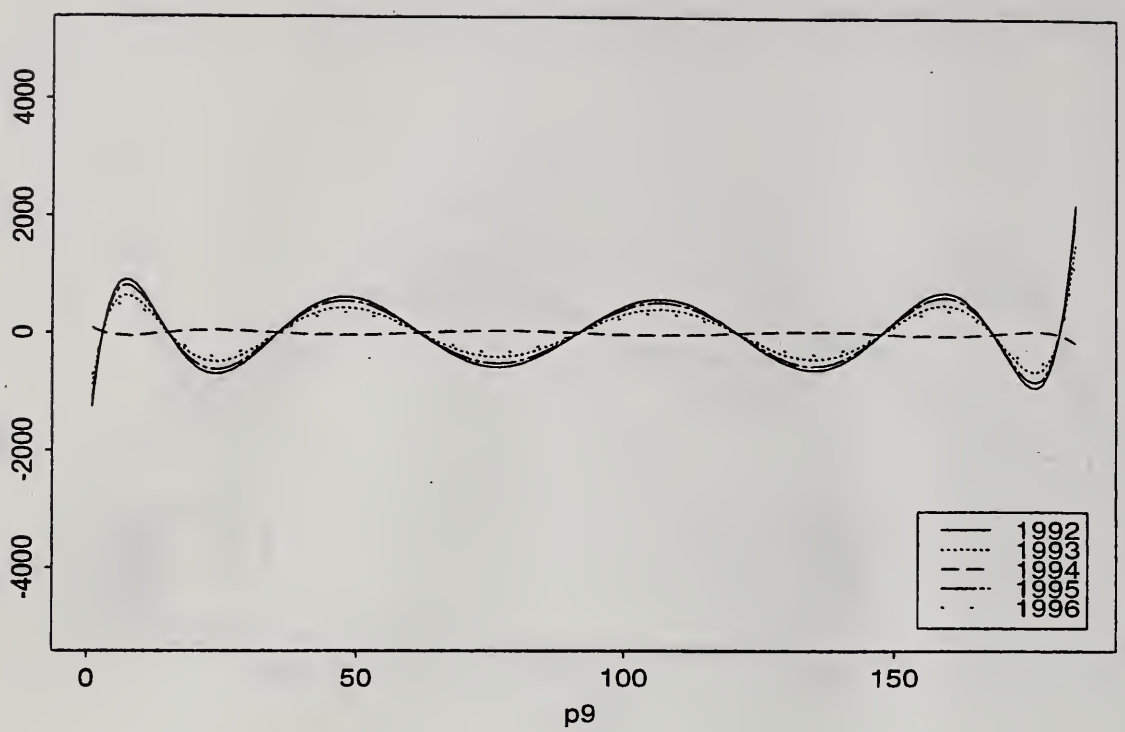

Figure 49: $\hat{\beta}_6 * p_9$

5.2 Integrating Rainfall

It has already been shown that rainfall plays an important role in predicting maximum daily power usage. One of the ways we have considered including this information is to use rainfall as a predictor for the residuals obtained under the polynomial fit. However, this method only improved the predictions of the drops in usage, not the peaks. Therefore, the benefit of using rainfall in this manner was not worth the required effort of predicting rainfall.

We also considered including rainfall as one of the predictors in the base model. However, the predictions under this modeling scheme were not as accurate, the model was more complicated, and the range of magnitude of each parameter was greater from one year to the next. At this point, we decided to postpone the inclusion of rainfall and move on to the next step, modeling the parameters.

6 Model Choices for the Parameters of the Base Model

We wanted to update the parameters of the base model as more information became available. The structural model framework allows the parameters to change with time and makes it easy to obtain predictions of the power usage for the remaining days based on any previous day. Therefore, we began modeling the parameters according to the structural model given below.

$$\begin{aligned} Y_t &= \mathrm{p}_t\boldsymbol{\beta}_t + W_t, \;\; W_t \sim \mathrm{N}(0, \sigma_w^2) \\ \boldsymbol{\beta}_{t+1} &= \boldsymbol{\beta}_t + \mathbf{V}_t, \;\; \mathbf{V}_t \sim \mathrm{N}(\mathbf{0}, Q) \\ t &= 1, \ldots, 183, \end{aligned}$$

where :

$$
\begin{aligned}
Y_t &= \text{the maximum daily power usage in kw} \\
\mathbf{p}_t &= \begin{bmatrix} p_0(t) & p_1(t) & p_2(t) & p_3(t) & p_5(t) & p_7(t) & p_9(t) \end{bmatrix} \\
\boldsymbol{\beta}_t &= \begin{bmatrix} \beta_0(t) & \beta_1(t) & \beta_2(t) & \beta_3(t) & \beta_4(t) & \beta_5(t) & \beta_6(t) \end{bmatrix}'
\end{aligned}
$$

with initial value

$$
\boldsymbol{\beta}_1 = \begin{bmatrix} 4036 & 1278 & -1488 & -1525 & 769 & -409 & 317 \end{bmatrix}'
$$

and

$$
Q = \begin{bmatrix}
\sigma_0^2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \sigma_1^2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \sigma_2^2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \sigma_3^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \sigma_4^2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \sigma_5^2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma_6^2
\end{bmatrix}
$$

Under this modeling scheme, the Kalman recursions simplify as follows:

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_{t+1} &= \hat{\boldsymbol{\beta}}_t + \Theta_t \Delta_t^{-1} (Y_t - \mathbf{p}_t \hat{\boldsymbol{\beta}}_t) \\
\Omega_{t+1} &= \Omega_t + Q - \Theta_t \Delta_t^{-1} \Theta_t'
\end{aligned}
$$

where

$$
\begin{aligned}
\Delta_t &= \mathbf{p}_t \Omega_t \mathbf{p}_t' + \sigma_w^2, \text{ and} \\
\Theta_t &= \Omega_t \mathbf{p}_t'.
\end{aligned}
$$

For the initial conditions, we let $\boldsymbol{\beta}_1$ equal the average of the parameter estimates obtained in Section 5.1 and $\Omega_1 = 0 \cdot I_{7.7}$. The best linear mean-squared h-step predictors $P_t Y_{t+h}$ are determined by the following equations:

$$
\begin{aligned}
P_t \boldsymbol{\beta}_{t+h} &= P_t \boldsymbol{\beta}_{t+1} = \hat{\boldsymbol{\beta}}_{t+1}, \text{ and} \\
P_t Y_{t+h} &= \mathbf{p}_{t+h} \hat{\boldsymbol{\beta}}_{t+1}.
\end{aligned}
$$

With this model, to obtain a prediction for day s (for example) based on the information available at time t ($s > t$), the interested party would simply calculate $\mathbf{p}_s \hat{\boldsymbol{\beta}}_{t+1}$. Therefore, the prediction for the remaining days based on the information available at time t would be a polynomial fit similar to the base model but with different coefficients.

The above model was fit using a Fortran program similar to the one used in Section 3.1. The function that the program seeks to optimize has many local minimums. Therefore, these estimates are not obtained from a function that has been minimized globally, hence they are not uniqe and may not be the true maximum likelihood estimates. However, reasonable variations of these estimates result in almost identical predictions. Thus, we did not feel it was necessary to spend the time required to find the exact maximum likelihood estimates if, in fact, they exist.

In our initial fit, we assumed that $Q = \sigma_Q^2 I_{7.7}$ ($\sigma_0^2 = \sigma_1^2 = \ldots = \sigma_6^2$). The maximum likelihood estimates of the parameters under this model are given in Table 6. Figures 50 through 54 depict the observed values ('o') and the one-step predictions ('*') for this model. Comparison of these figures with those resulting from earlier models shows the gain in predictive power achieved using the base model approach.

Next we fit the same theoretical model but allowed the variances of the β's to differ ($\sigma_0^2 \neq \sigma_1^2 \neq \ldots \neq \sigma_6^2$). Removal of the equal variances constraint resulted in $\hat{\sigma}_1^2$ through $\hat{\sigma}_6^2$ being close to zero and $\hat{\sigma}_0^2$ being quite large. Therefore, the one-step predictions were essentially found by moving the base model polynomial up or down so that it passed through the observed data. This resulted in very poor h-step predictions ($h > 1$). For this reason, we kept the constraint of equal variances in our model framework. The interest at this point was in improving the predictions obtained under this model.

Table 6: Maximum Likelihood Estimates

Parameter	1992	1993	1994	1995	1996
σ_w^2	29,108	86,140	45,242	41,222	230,771
σ_Q^2	136,514	142,993	125,771	77,909	212,309

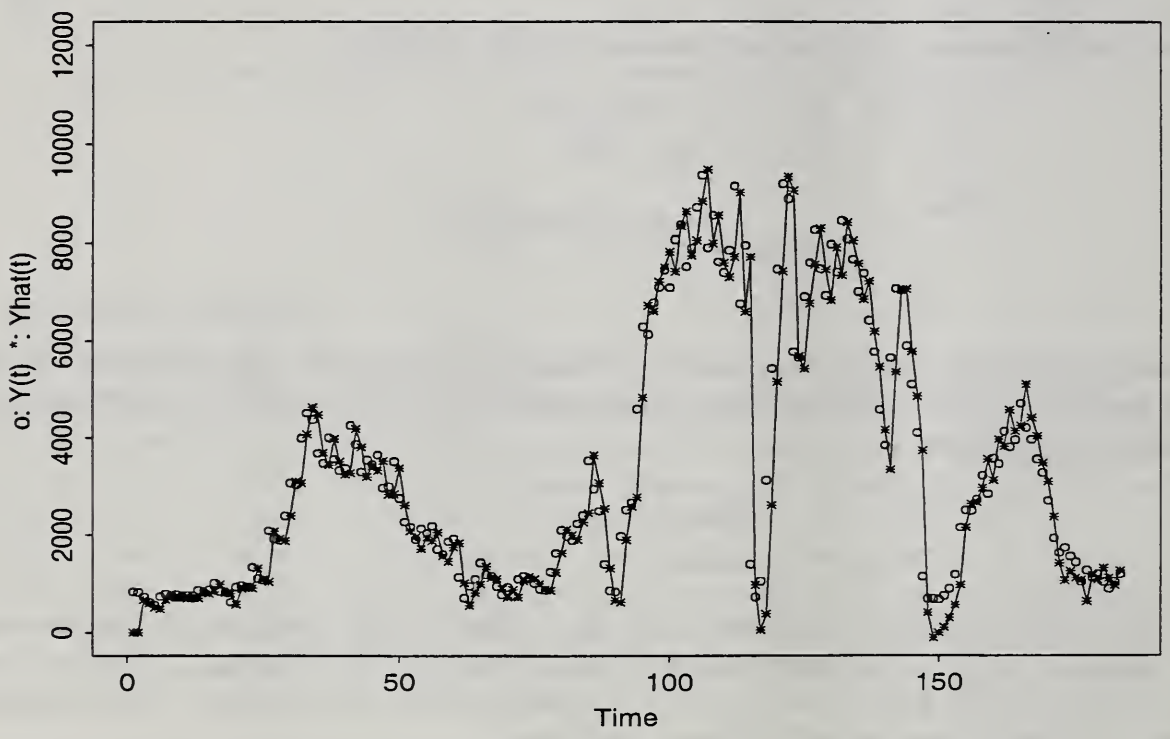

Figure 50: One-step Predictions from Structural Model - 1992

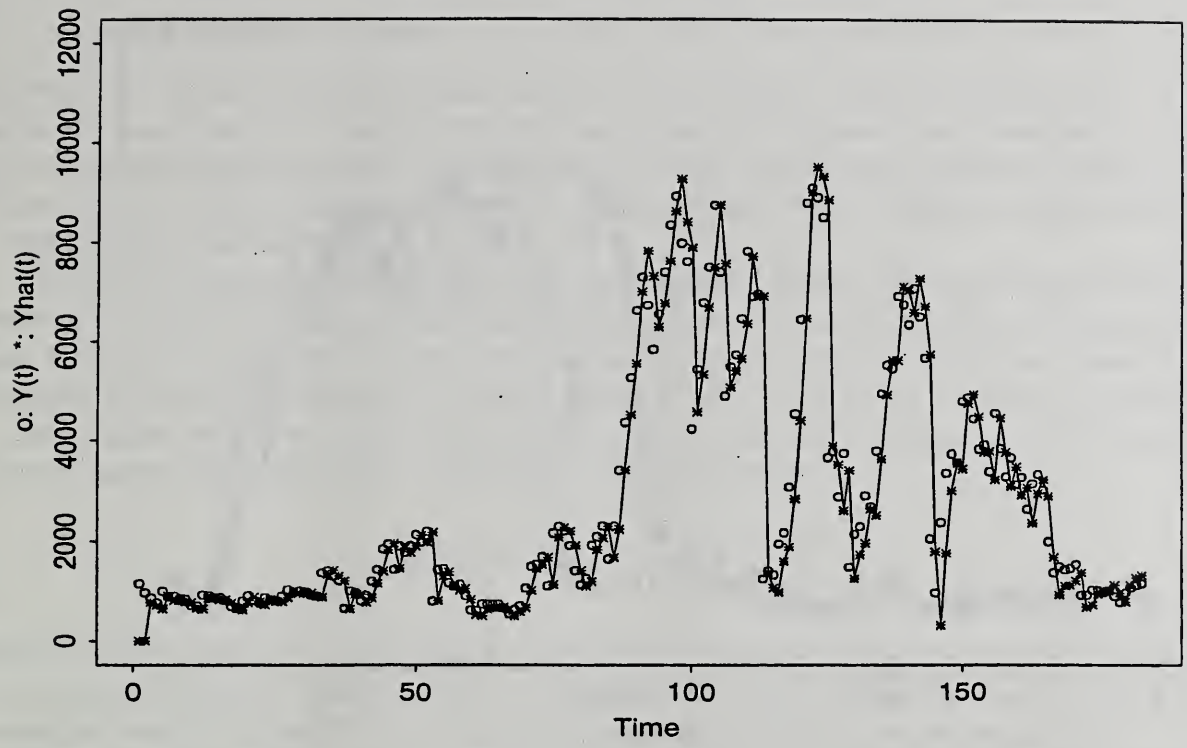

Figure 51: One-step Predictions from Structural Model - 1993

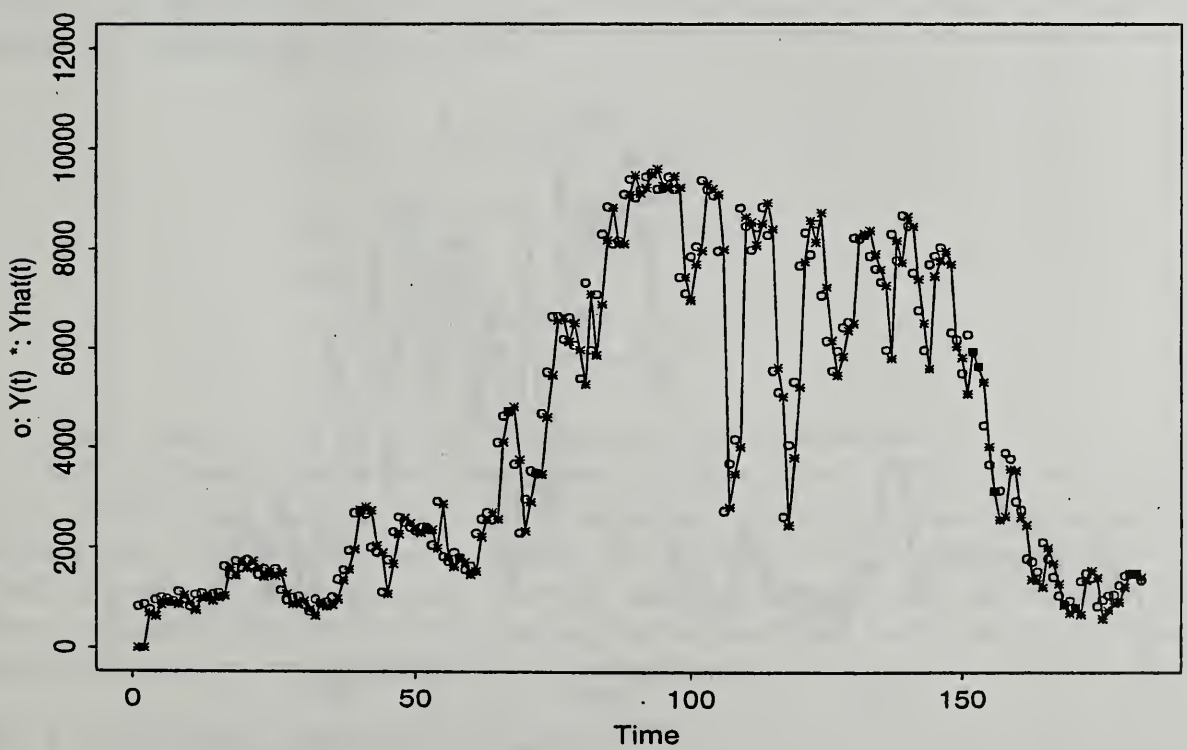

Figure 52: One-step Predictions from Structural Model - 1994

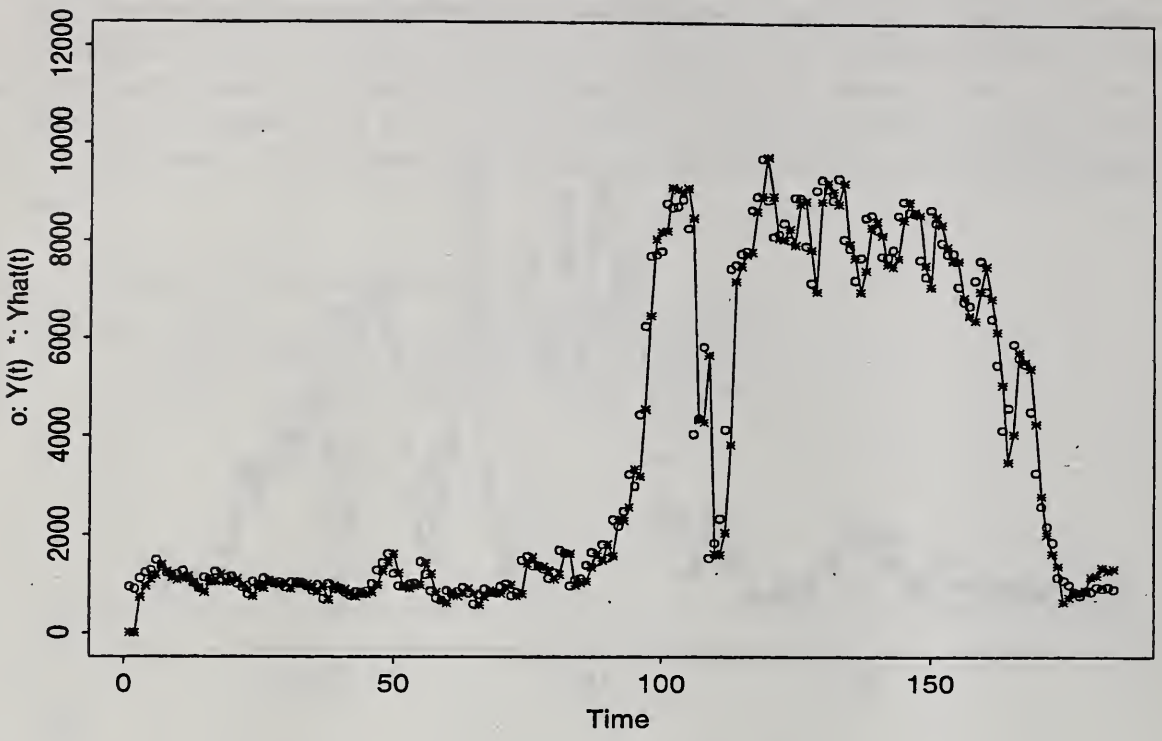

Figure 53: One-step Predictions from Structural Model - 1995

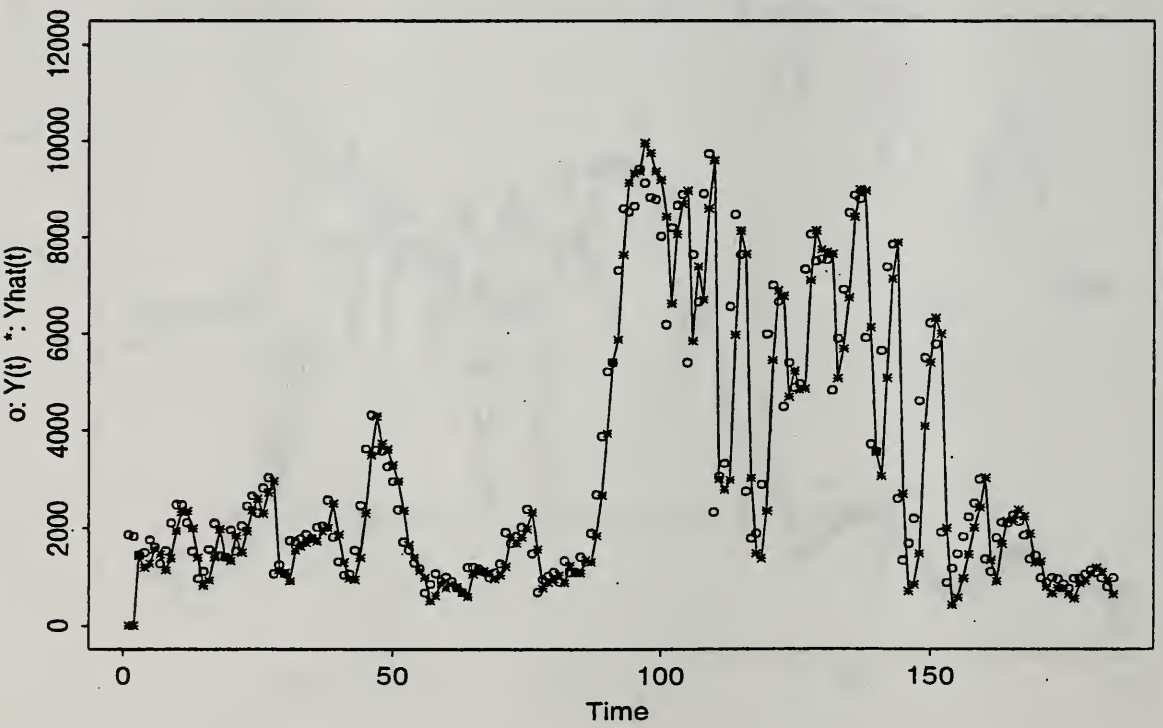

Figure 54: One-step Predictions from Structural Model - 1996

7 Improving the Parameter Modeling Strategy

7.1 Robust Kalman Filter

Although the parameters of the base model were fit using a robust technique, the "updated" parameter estimates, $\hat{\beta}_t$, from Section 6 obtained using the Kalman recursions were not robust. This is because the Kalman equations give the best linear mean-square predictors which are heavily influenced by outlying observations. At this point we began researching alternative ways to arrive at robust parameter estimates.

Our search resulted in the discovery of a paper entitled, *Robust Kalman Filter and Its Application in Time Series Analysis,* by Tomáš Cipra and Rosario Romera (Cipra and Romera [6]). This paper suggested converting the structural model to a linear regression model and then minimizing a suitable robustifying function in order to obtain a robust form of the Kalman recursions.

Using these results, we began searching for a suitable robustifying function. We first tried what is known as the Huber function. This function is given below.

$$\psi_H(z) = \begin{cases} z, & for\ |z| \le c, \\ c\mathrm{sgn}(z), & for\ |z| > c, \end{cases}$$

where c is some constant. (The reader is referred to Lehmann [9] or Huber [8] for further information on the Huber function.) This function did not provide us with satisfactory predictions because it gives equal weight to positive and negative residuals that are of the same magnitude.

We eventually decided upon an asymmetric squared error function which weights the positive residuals ten times greater than the negative ones. The number ten is somewhat arbitrary. As long as the positive residuals receive significantly more weight, the fit will be essentially the same. Applying the results of the paper by Cipra and Romera to this function, the robust recursive formulas for obtaining $\hat{\beta}_{t+1}$ are as follows:

$$\begin{aligned} \hat{\beta}_{t+1} &= \hat{\beta}_t + \frac{\Omega_t \mathbf{p}_t'}{\mathbf{p}_t \Omega_t \mathbf{p}_t' + \sigma_w^2/c}(Y_t - \mathbf{p}_t\hat{\beta}_t) \\ \Omega_{t+1} &= \Omega_t + Q - \frac{\Theta_t \Theta_t'}{\mathbf{p}_t \Omega_t \mathbf{p}_t' + \sigma_w^2/c}, \end{aligned}$$

where

$$c = \begin{cases} 10, & \text{if } (Y_t - \hat{Y}_t) > 0, \\ 1, & \text{otherwise.} \end{cases}$$

For the case when $Y_t < \hat{Y}_t$, we obtain the same predictors as in Section 6.

In order to simplify the optimization process, we chose $\sigma_w^2 = 10{,}000$ and $\sigma_Q^2 = 100$ (recall $Q = \sigma_Q^2 I_{7\cdot7}$). Again, these numbers were chosen somewhat arbitrarily. Although several optimization programs were used to find these estimates, the function has many local minimums making it difficult to arrive at a global minimum. However, the fit changed very little for reasonable variations of these estimates. This was beneficial not only for the reason just mentioned, but also because we were able to use the same estimates for all five years.

7.2 Incorporating Weather Variables

We began including weather variables in our modeling framework at the same time as our search for more robust techniques was initiated. Rainfall was the first weather variable considered. We knew that rainfall was responsible for the extreme drops in usage and that large amounts of precipitation

had a prolonged affect in many cases. We wanted to use this information to prevent the parameter estimates from changing drastically in these cases, not to obtain more accurate predictions of the drops. Therefore, we allowed the variance of the polynomial coefficients (σ_Q^2) to change according to rainfall. This was done by first creating a function to account for the additive affect of rainfall. After much debate and experimentation, we settled on the function $h(t) = 0.6R(t) + 0.4h(t-1)$, where R(t) was the rainfall for day t. We let $\sigma_Q^2(t) = 100\exp^{-3.014h(t)}$. This technique allows us to include the effects of rainfall in our one-step predictions without having to predict precipitation. For our h-step predictions ($h > 1$), we predicted rainfall to be $0mm$ with satisfying results. The value 3.014 was found by averaging the optimum values of a in the function $100\exp^{-a\cdot h(t)}$ for all five available years. The optimum values of a were found using the IMSL routine DUVMIF and are given in Table 7.2.

Table 7.2: Optimum values for a

	1992	1993	1994	1995	1996	average
a	4.475	-0.567	2.049	1.719	7.404	3.014

In addition to including rainfall, we investigated the inclusion of ETAcorn. We began by adding it to our model in the same manner as we had for precipitation. However, this did not significantly improve our predictions. We then included ETAcorn as a parameter in the base model. Although this requires the user to have an estimate of ETAcorn for the entire year, it can be calculated at the beginning of the year assuming normal conditions. Including ETAcorn in this manner did improve our predictions. The resulting model (as given below) is a variation of that given in Section 6.

$$
\begin{aligned}
Y_t &= \mathbf{p}_t\boldsymbol{\beta}_t + W_t, \quad W_t \sim \mathrm{N}(0, 10,000) \\
\boldsymbol{\beta}_{t+1} &= \boldsymbol{\beta}_t + \mathbf{V}_t, \quad \mathbf{V}_t \sim \mathrm{N}(\mathbf{0}, \sigma_Q^2(t)\cdot I_{8\cdot 8}) \\
t &= 1, \ldots, 183, \\
&\text{where :} \\
Y_t &= \text{the maximum daily power usage in kw} \\
\mathbf{p}_t &= \left[\begin{array}{cccccccc} p_0(t) & p_1(t) & p_2(t) & p_3(t) & p_5(t) & p_7(t) & p_9(t) & eta(t) \end{array}\right] \\
\boldsymbol{\beta}_t &= \left[\begin{array}{cccccccc} \beta_0(t) & \beta_1(t) & \beta_2(t) & \beta_3(t) & \beta_4(t) & \beta_5(t) & \beta_6(t) & \beta_7(t) \end{array}\right]' \\
&\text{with initial value} \\
\boldsymbol{\beta}_1 &= \left[\begin{array}{cccccccc} 3793 & 1083 & -1167 & -1344 & 637 & -318 & 245 & 268 \end{array}\right]' \\
\sigma_Q^2(t) &= 100\exp^{-3.014\cdot h(t)} \\
&\text{and} \\
h(t) &= 0.6R(t) + 0.4h(t-1).
\end{aligned}
$$

For the initial conditions, we let Ω_1 equal an 8 by 8 matrix of zeros and we chose β_1 to be the average of the parameter estimates obtained in a similar manner to those found in Section 5.1. The asymmetric loss function used in obtaining β_1 for this model was ten times the squared-error loss for positive residuals and squared-error loss for negative residuals. The h-step predictions are found in the same manner as that described in Section 6. Namely, the prediction of maximum power usage for day s based on the information available up to time t ($t < s$) is found by calculating $\mathbf{p}_s\hat{\boldsymbol{\beta}}_{t+1}$.

Five sets of plots depicting the observed values ('o') and h-step predictions ('*'), $h = 1, 3, 5, 7$, for all years under this model are included below as Figures 55-74 (Figures 55-59 give the one-step predictors, Figures 60-64 the three-step predictors, Figures 65-69 the five-step predictors and Figures 70-74 the seven-step predictors). The plotted predicted values ('*') are actually the maximum of zero and the values predicted by the model. The plots of the one-step predictors include darkened circles indicating that there was a positive amount of precipitation on the previous day. As was discussed earlier, the polynomial coefficients will change by a negligible amount when there has been a significant amount of rainfall; thus the predictions will closely follow a polynomial curve for the next several days after a large rainfall. This effect due to rainfall can be seen by careful examination of Figures 55-59. Comparison of the one-step predictions from this model with predictions from earlier models shows the gain in predictive power obtained by use of this modeling scheme. Even the predictions made seven days in advance are reasonably good when one considers that they are based only on historical data, past rainfall and ETAcorn. Residual plots for the one-step and seven-step predictors are given in Figures 75-84. Residuals are calculated as the observed value minus the predicted value.

The Fortran program used to fit the model described in this section is similar to the programs used in Sections 3.1 and 6 with adjustments made to incorporate the changes discussed above. In addition, the use of Hooke and Jeeve's algorithm to find the maximum likelihood estimates of σ_w^2 and Q is no longer required since, for this model, σ_w^2 is fixed and σ_Q^2 is determined by the function $h(t)$ previously defined.

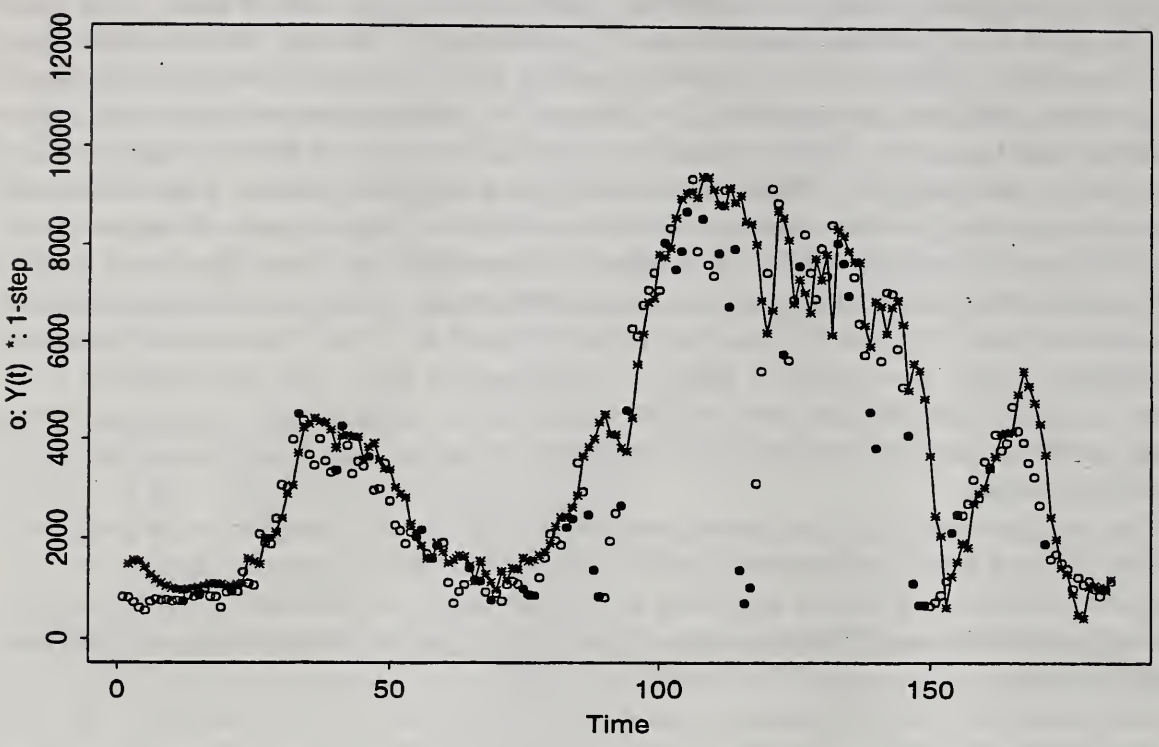

Figure 55: One-step Predictions from Final Model - 1992

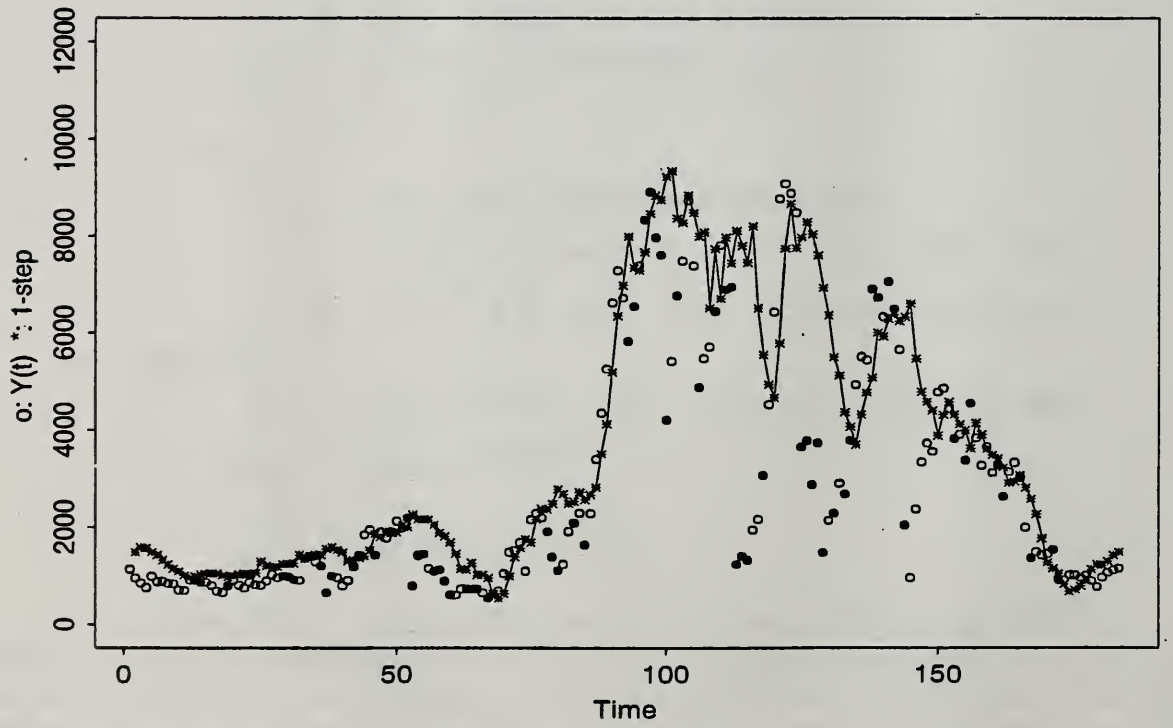

Figure 56: One-step Predictions from Final Model - 1993

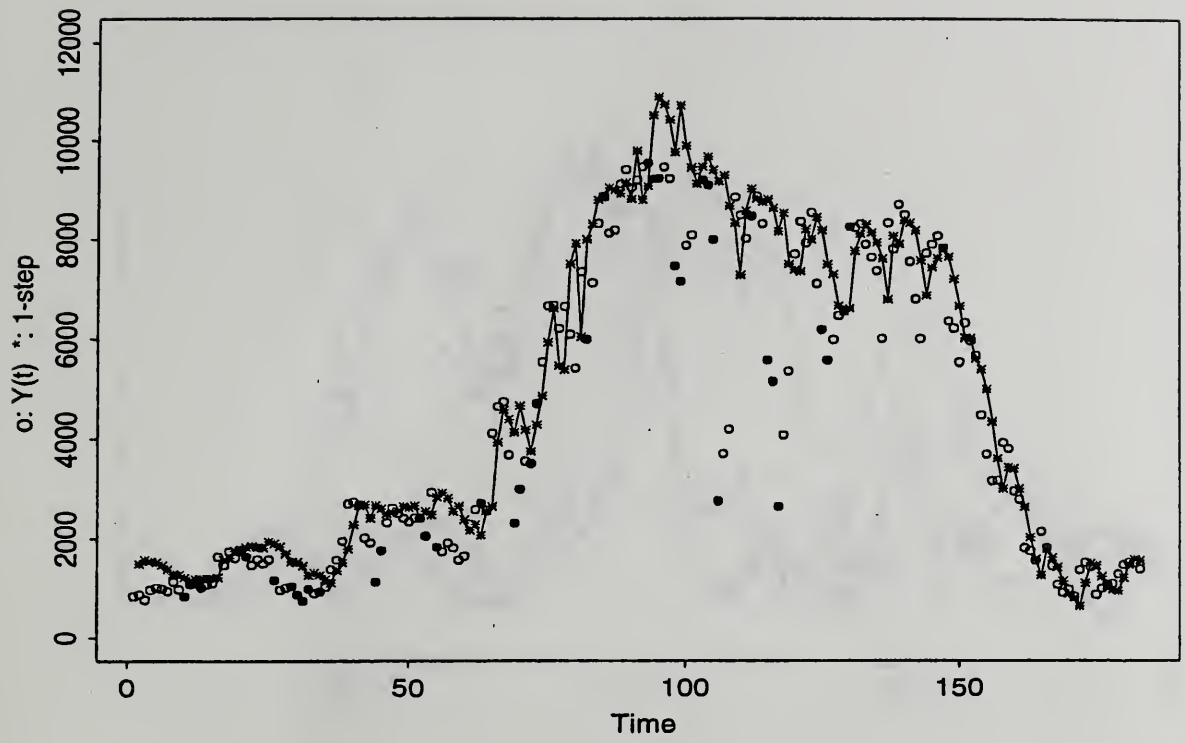

Figure 57: One-step Predictions from Final Model - 1994

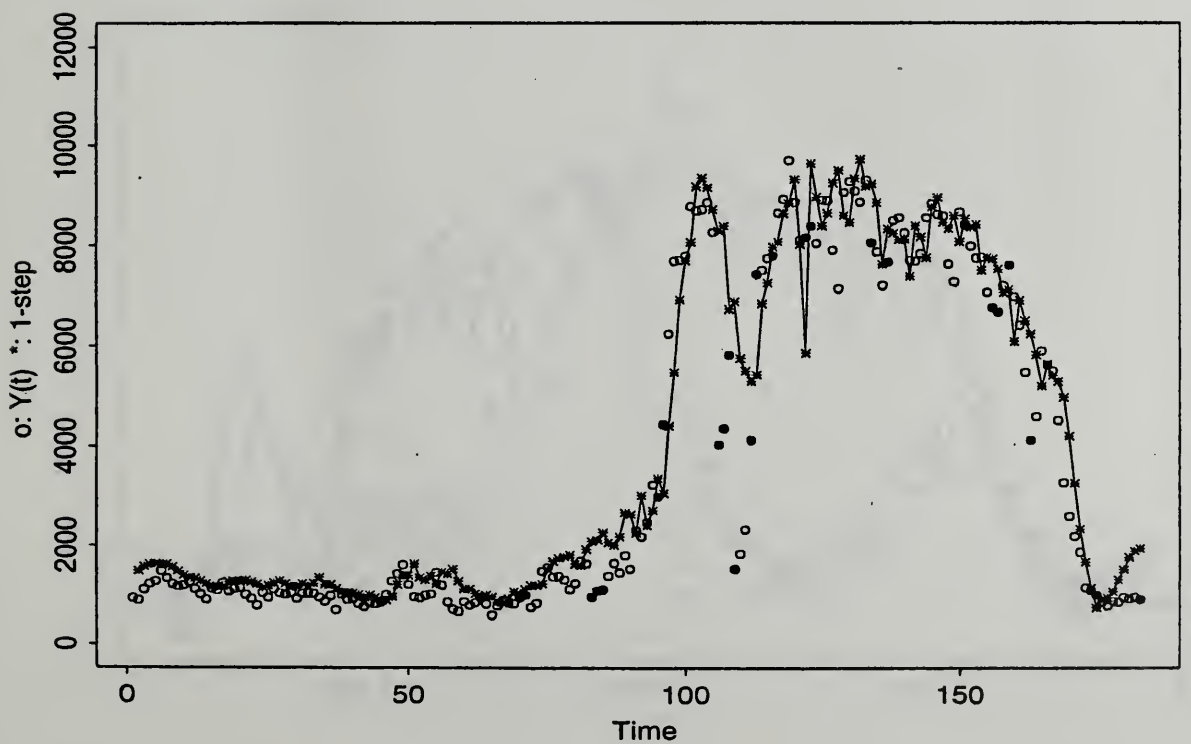

Figure 58: One-step Predictions from Final Model - 1995

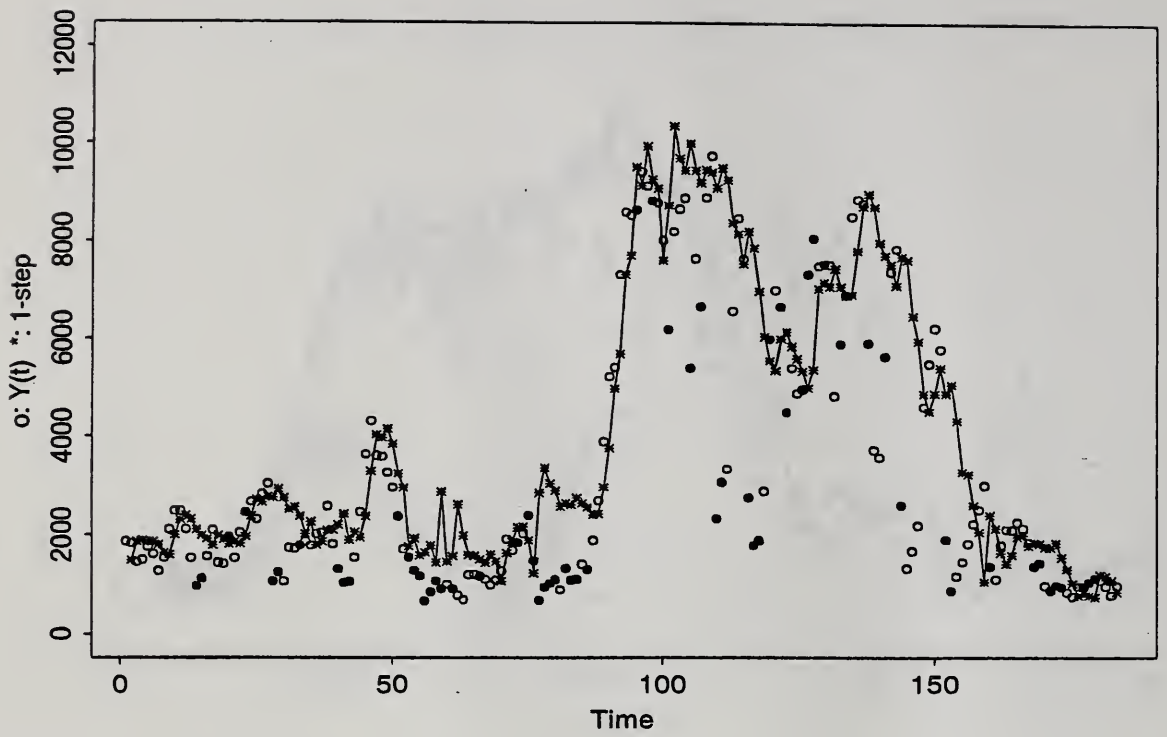

Figure 59: One-step Predictions from Final Model - 1996

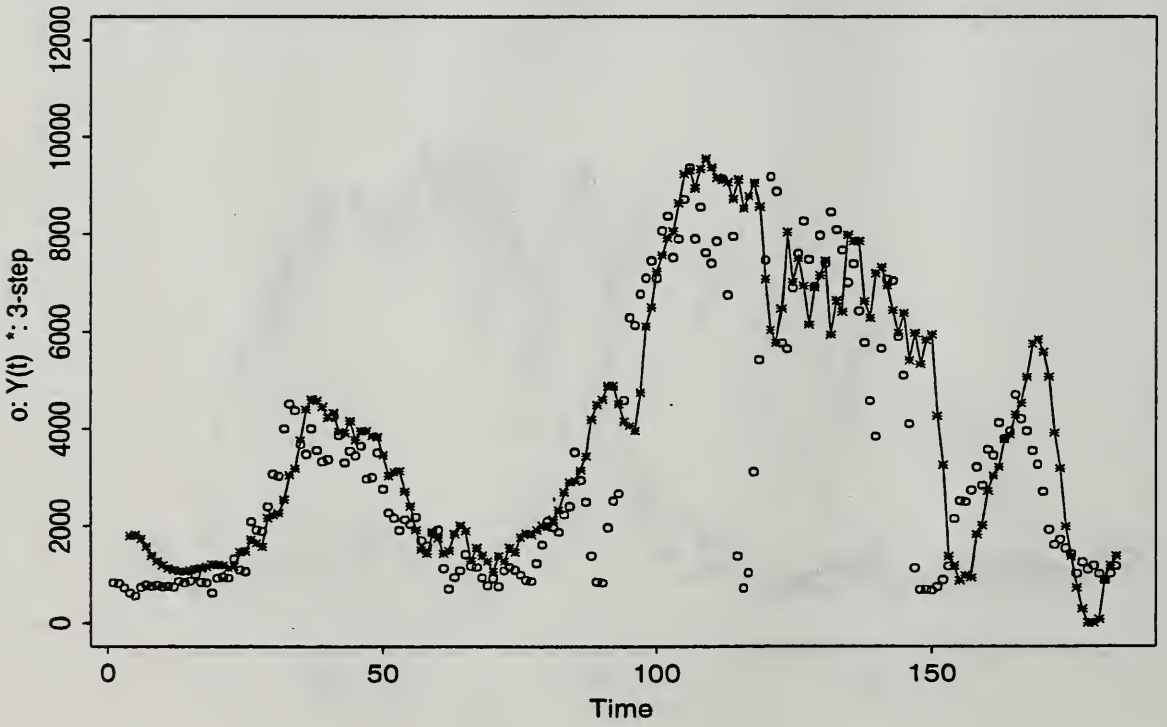

Figure 60: Three-step Predictions from Final Model - 1992

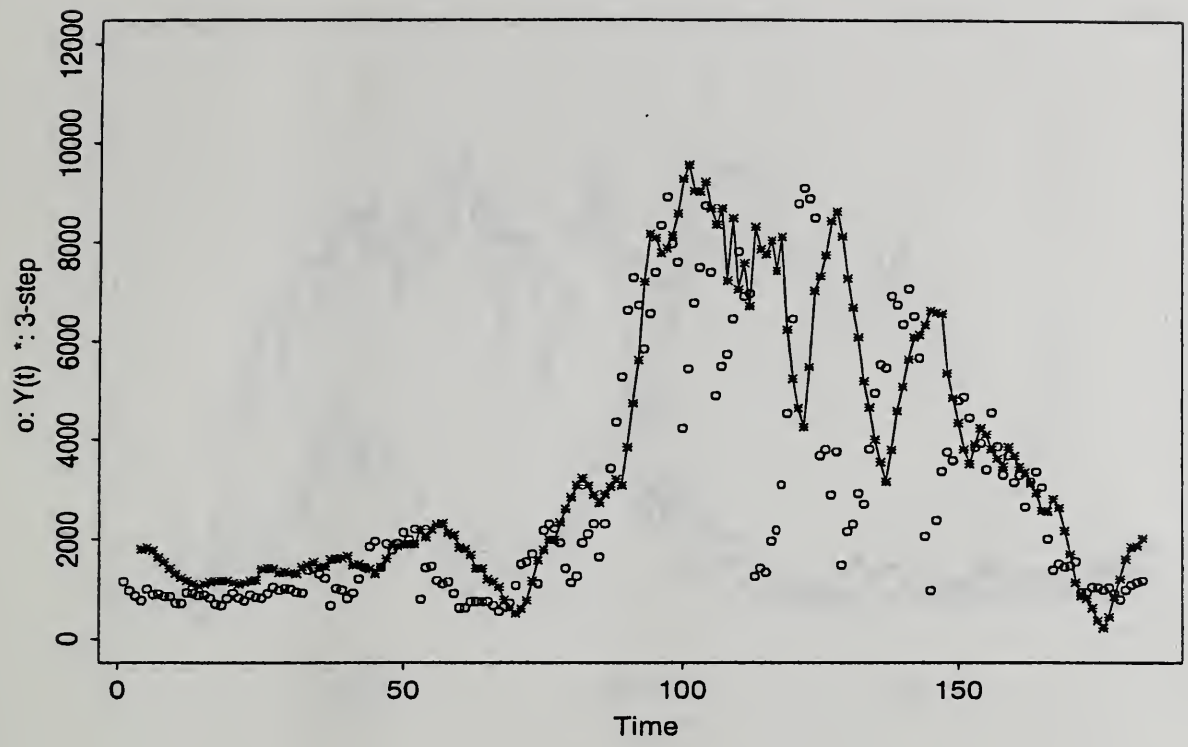

Figure 61: Three-step Predictions from Final Model - 1993

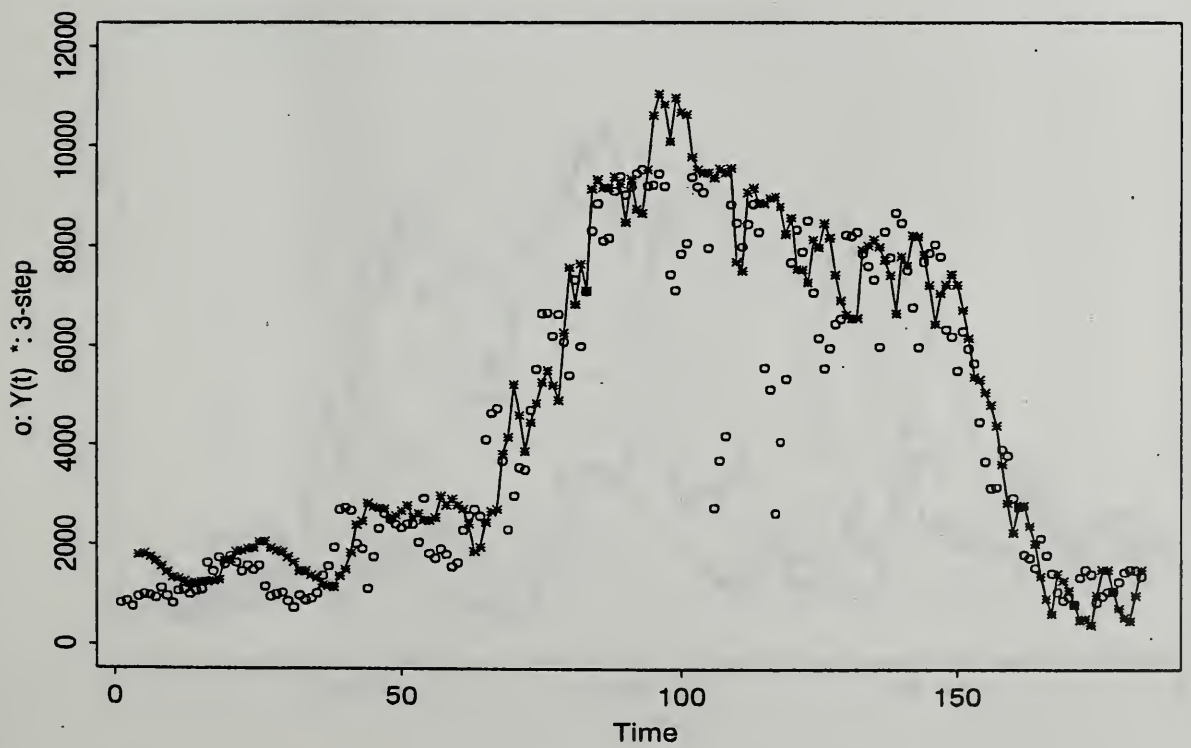

Figure 62: Three-step Predictions from Final Model - 1994

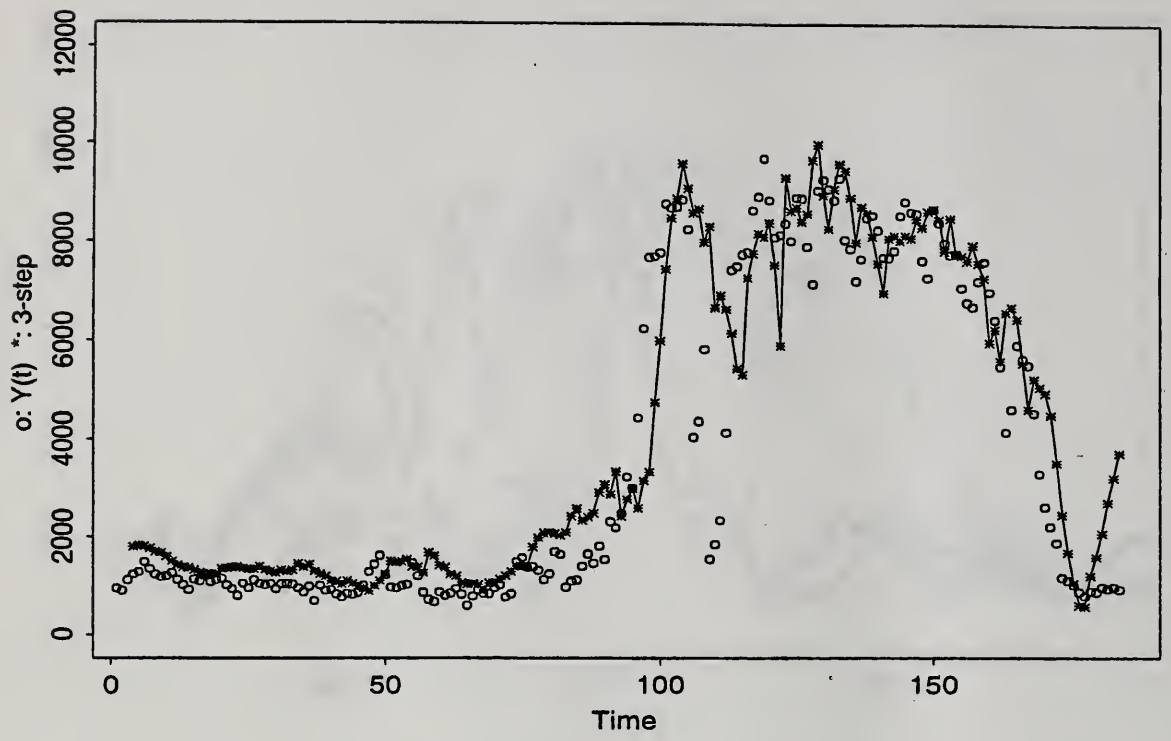

Figure 63: Three-step Predictions from Final Model - 1995

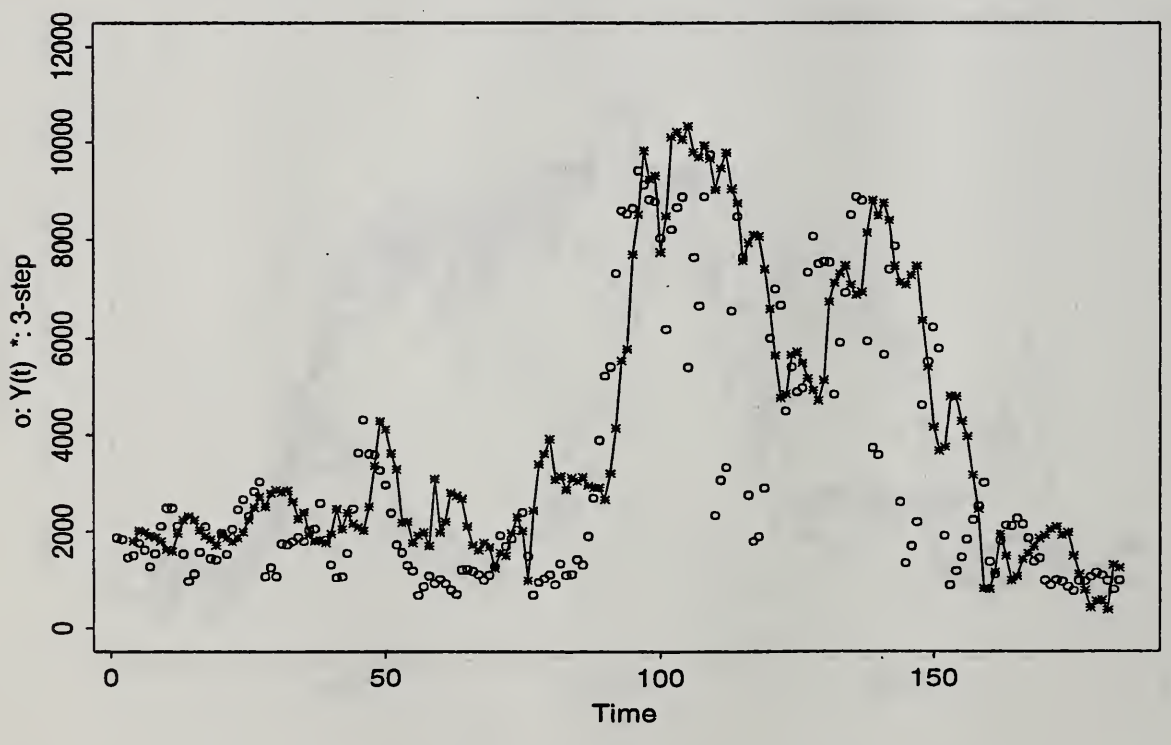

Figure 64: Three-step Predictions from Final Model - 1996

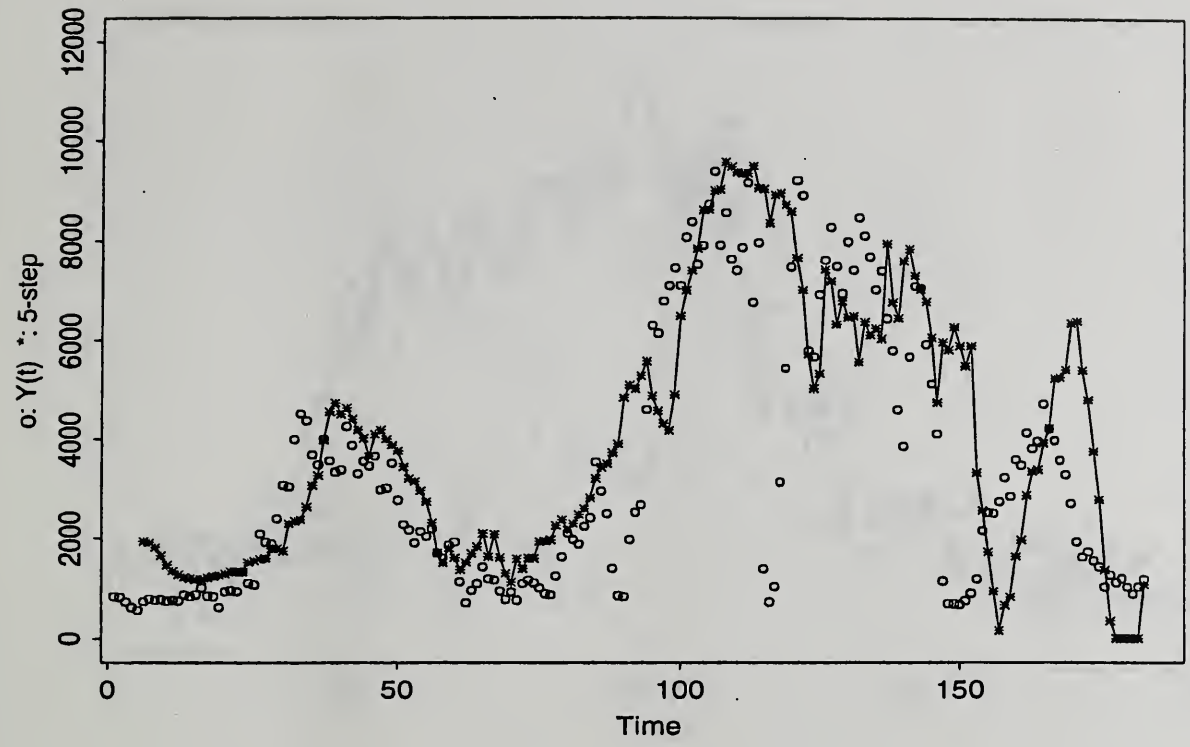

Figure 65: Five-step Predictions from Final Model - 1992

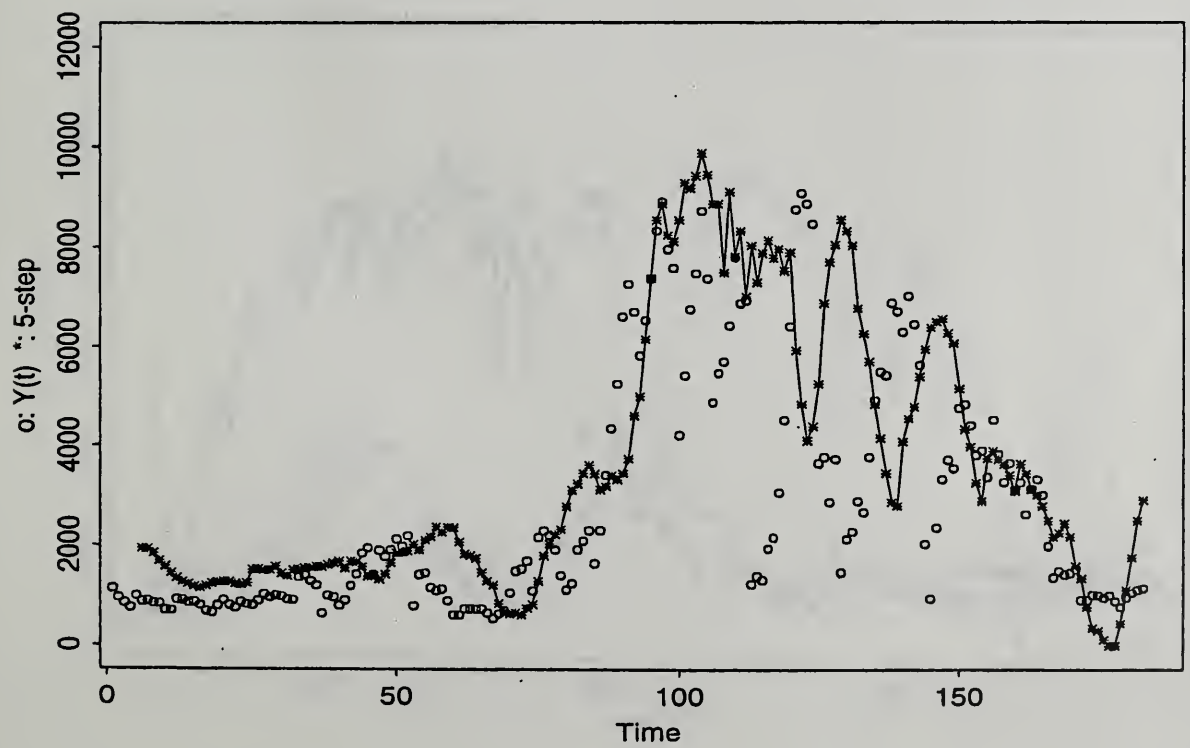

Figure 66: Five-step Predictions from Final Model - 1993

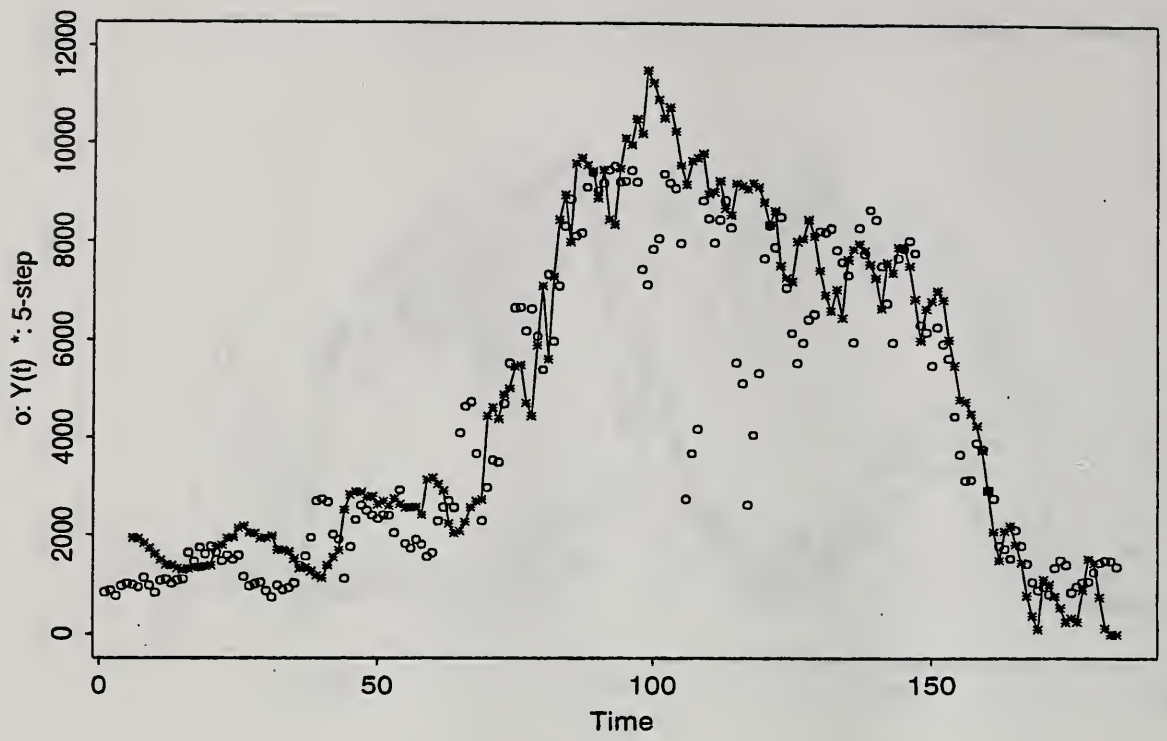

Figure 67: Five-step Predictions from Final Model - 1994

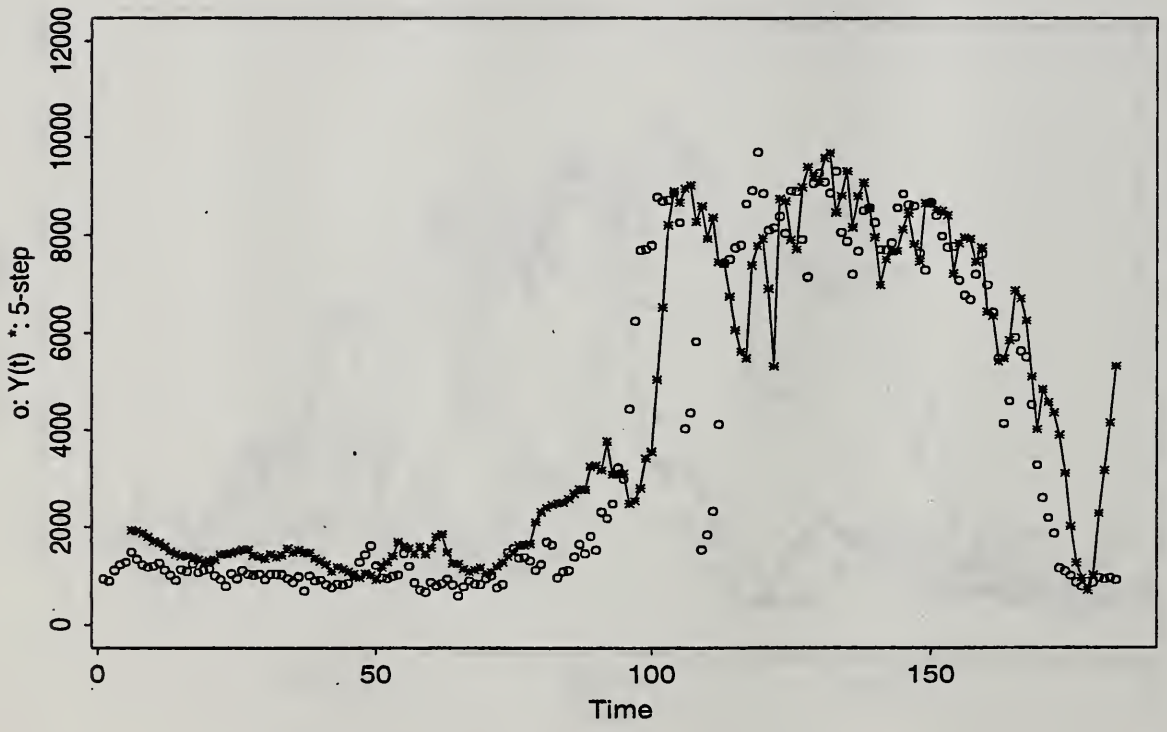

Figure 68: Five-step Predictions from Final Model - 1995

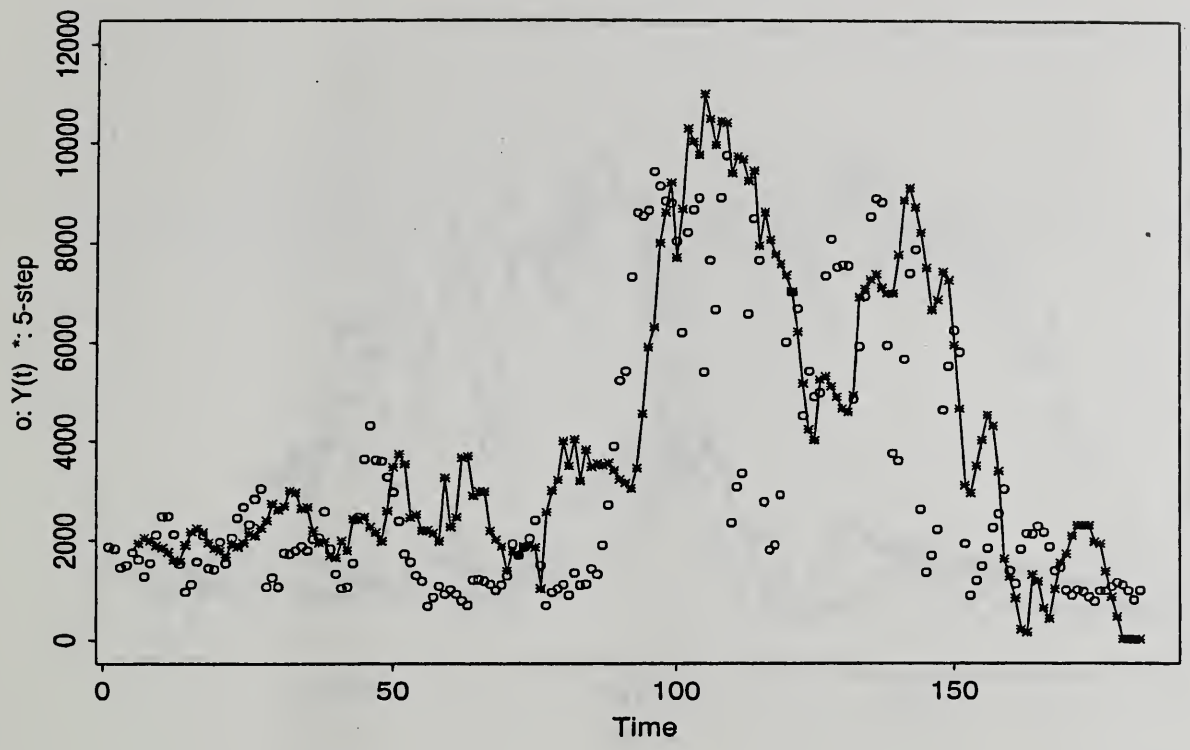

Figure 69: Five-step Predictions from Final Model - 1996

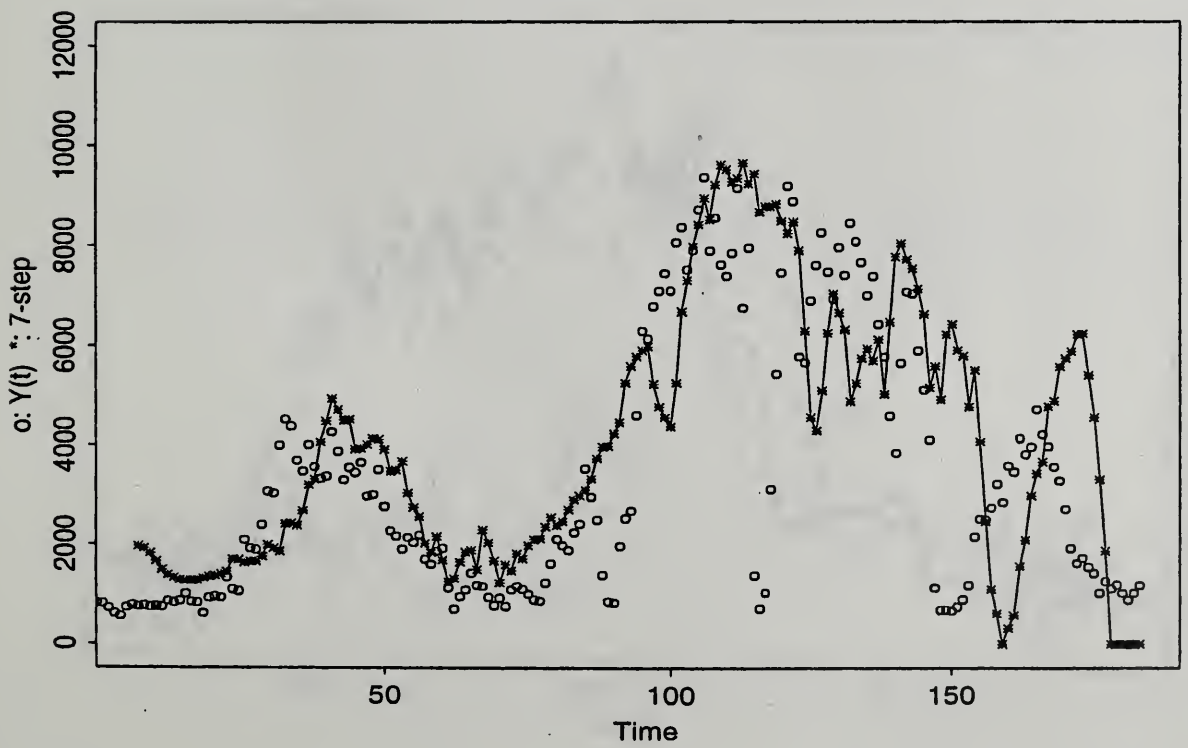

Figure 70: Seven-step Predictions from Final Model - 1992

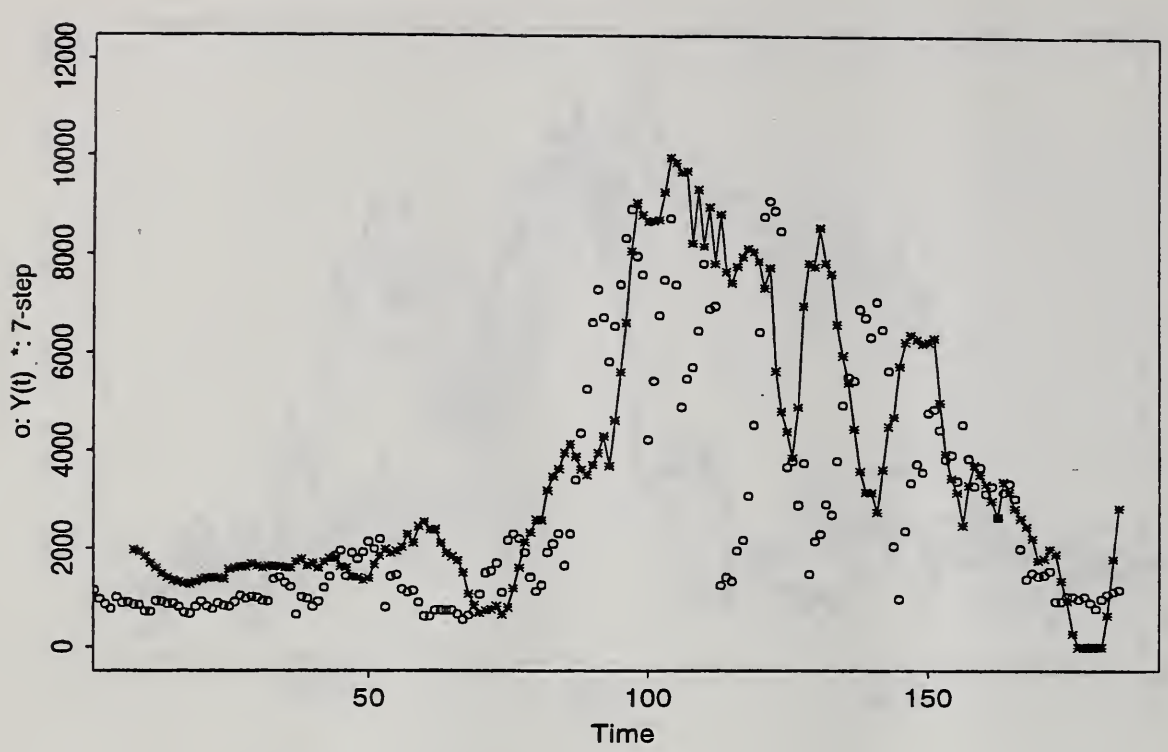

Figure 71: Seven-step Predictions from Final Model - 1993

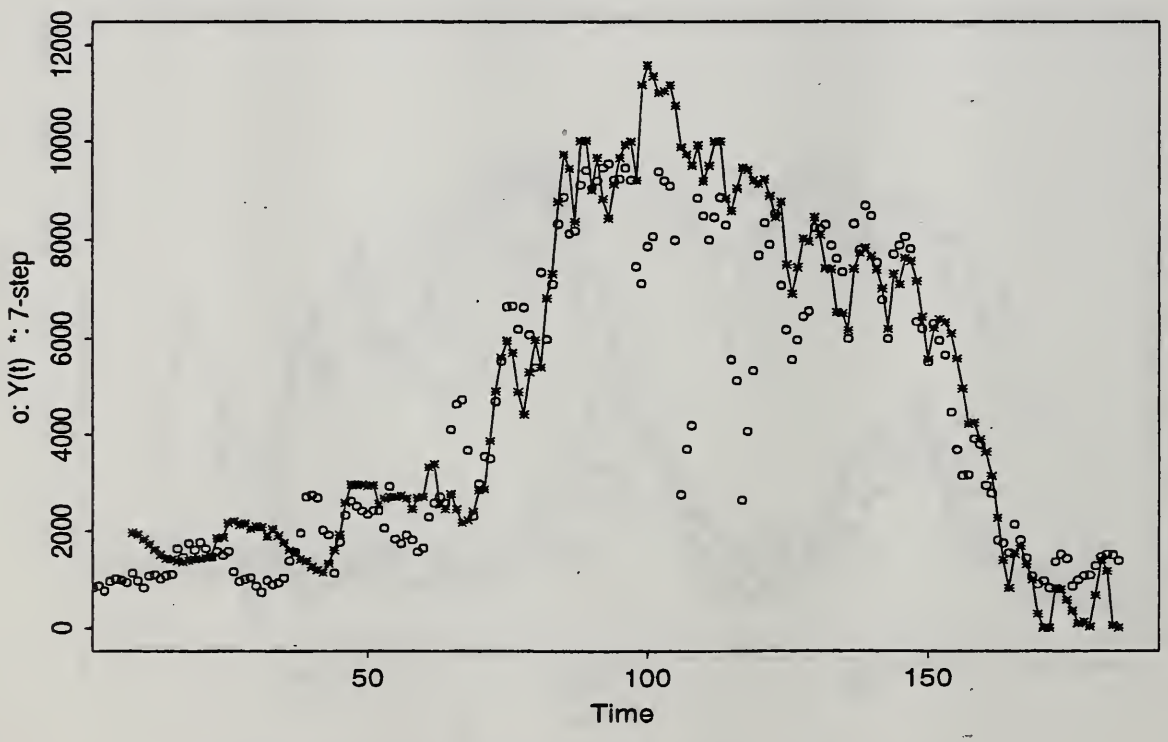

Figure 72: Seven-step Predictions from Final Model - 1994

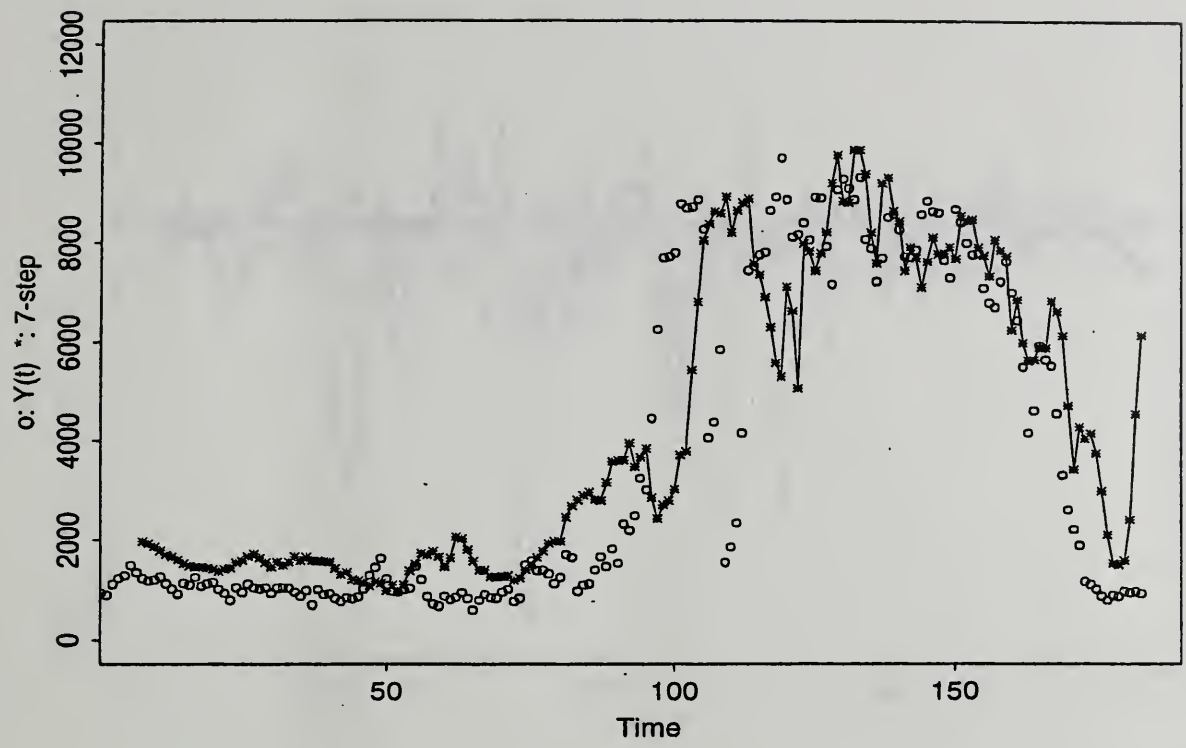

Figure 73: Seven-step Predictions from Final Model - 1995

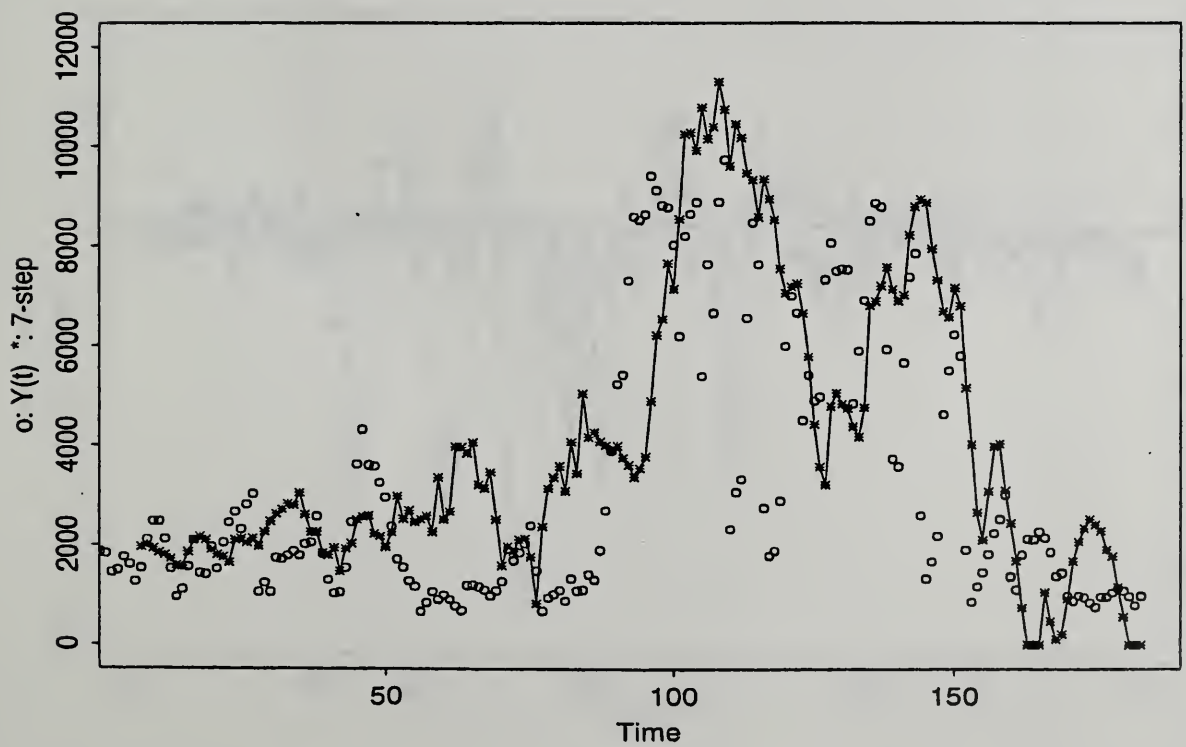

Figure 74: Seven-step Predictions from Final Model - 1996

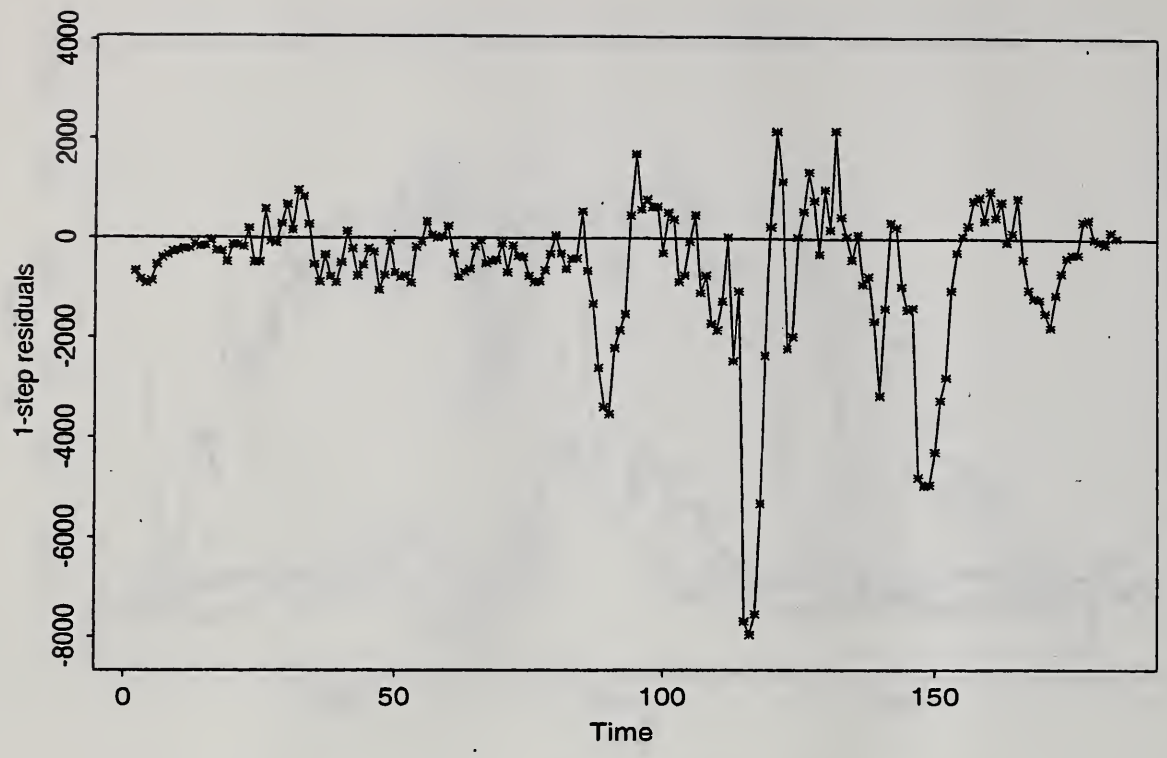

Figure 75: Residuals from One-step Predictions - 1992

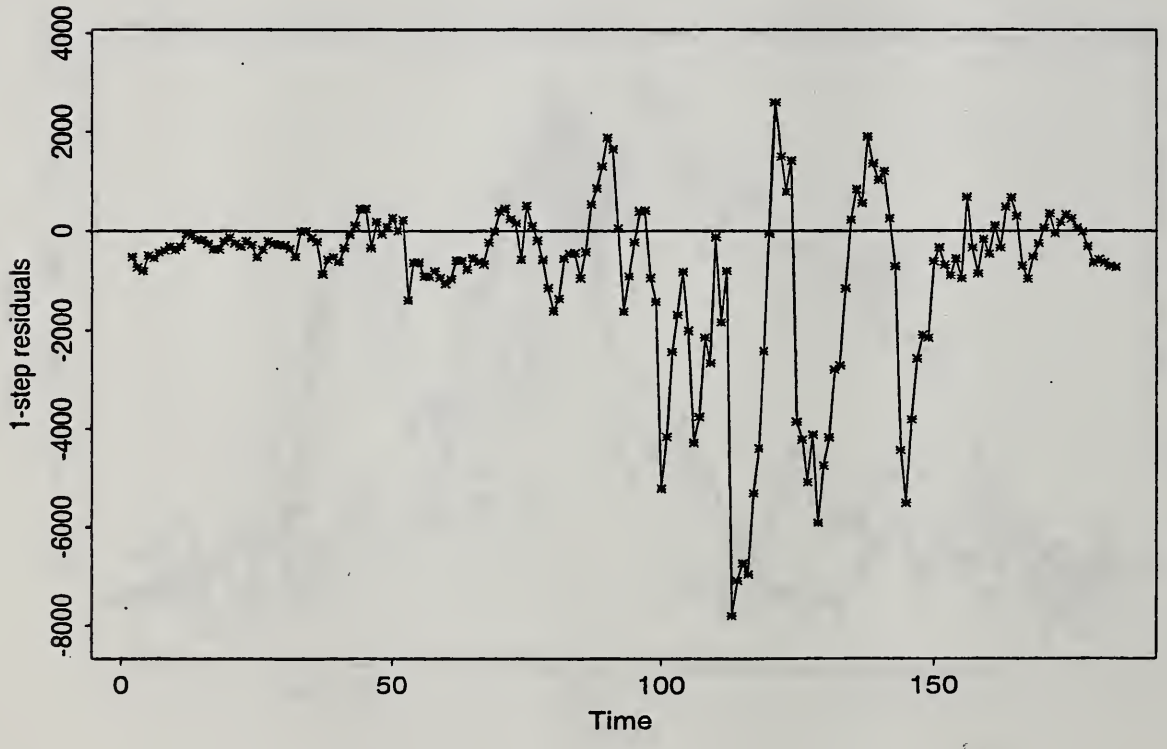

Figure 76: Residuals from One-step Predictions - 1993

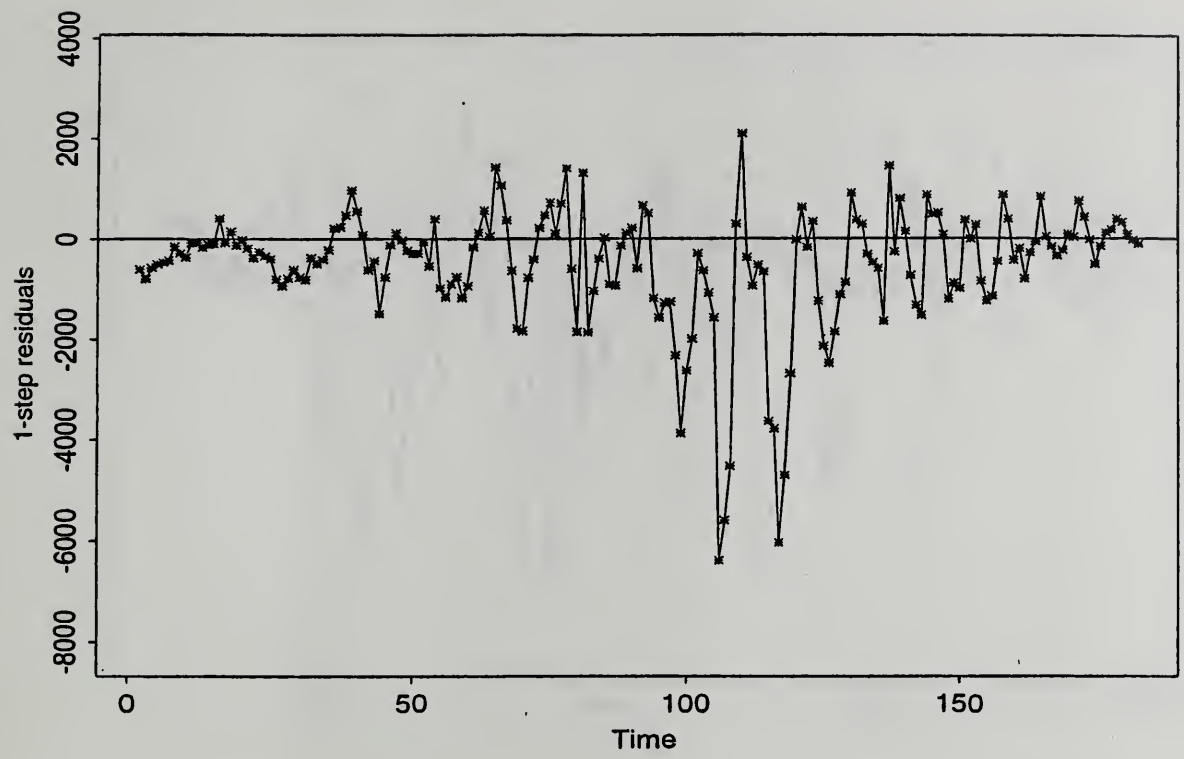

Figure 77: Residuals from One-step Predictions - 1994

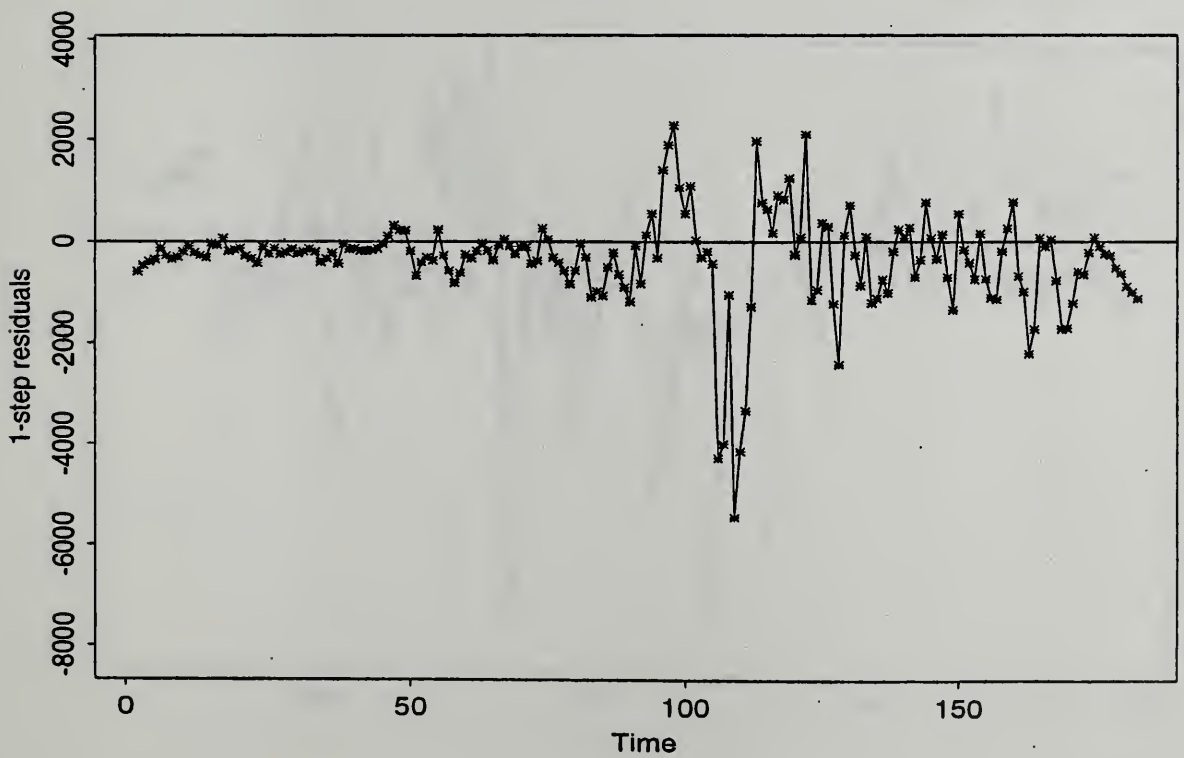

Figure 78: Residuals from One-step Predictions - 1995

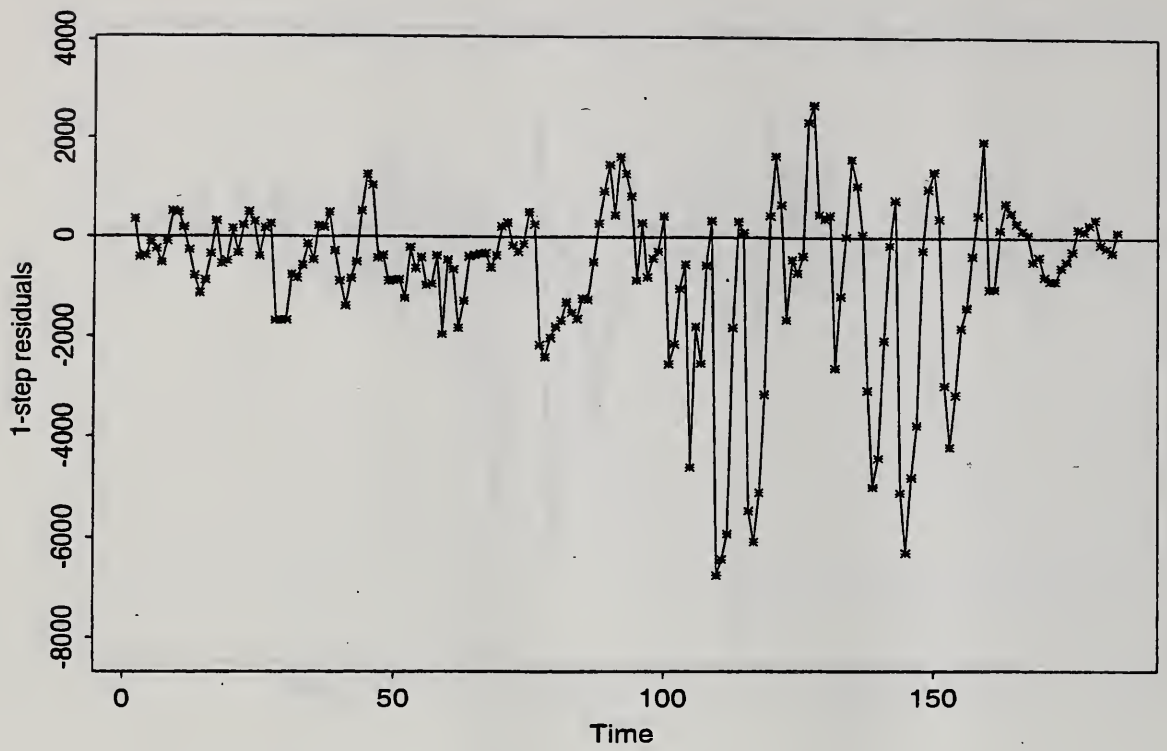

Figure 79: Residuals from One-step Predictions - 1996

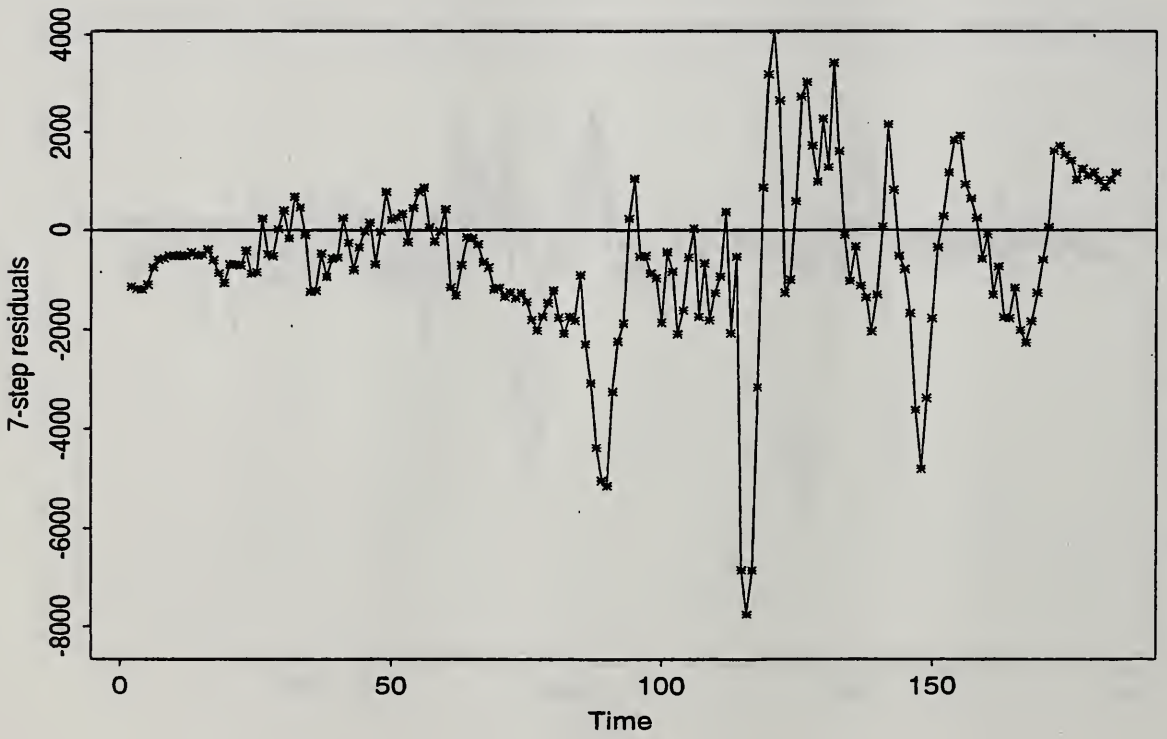

Figure 80: Residuals from Seven-step Predictions - 1992

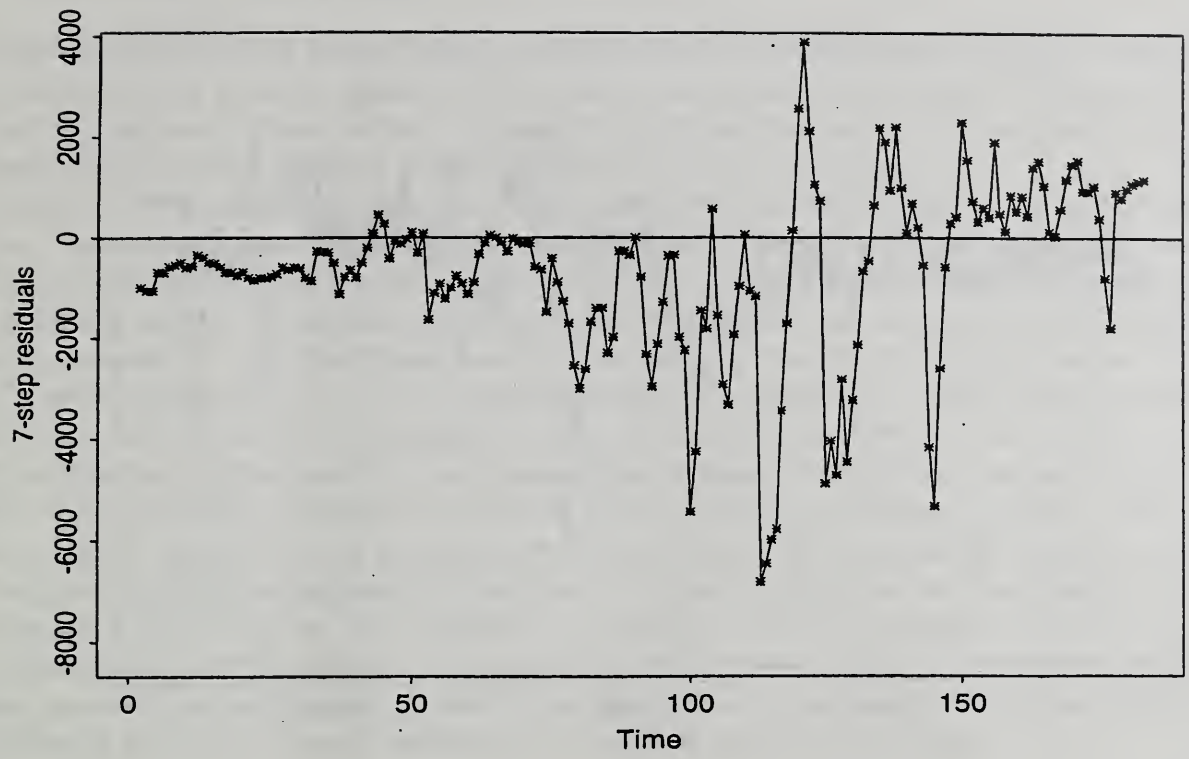

Figure 81: Residuals from Seven-step Predictions - 1993

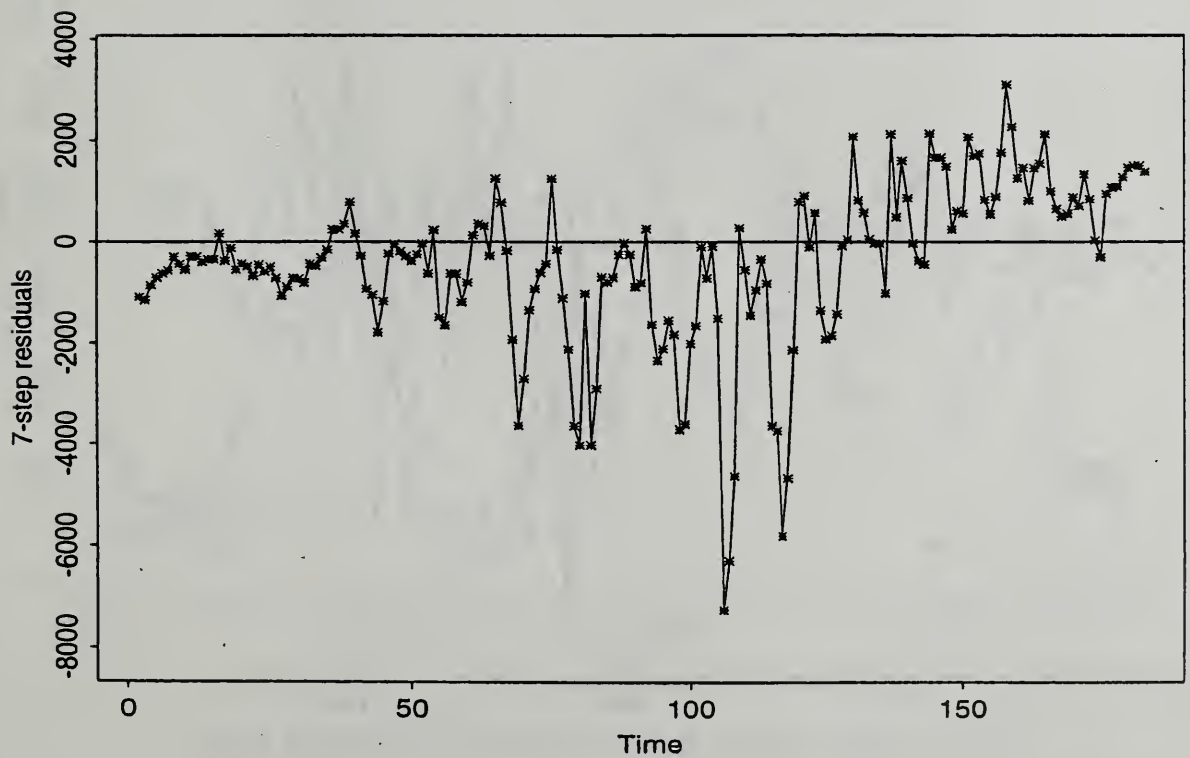

Figure 82: Residuals from Seven-step Predictions - 1994

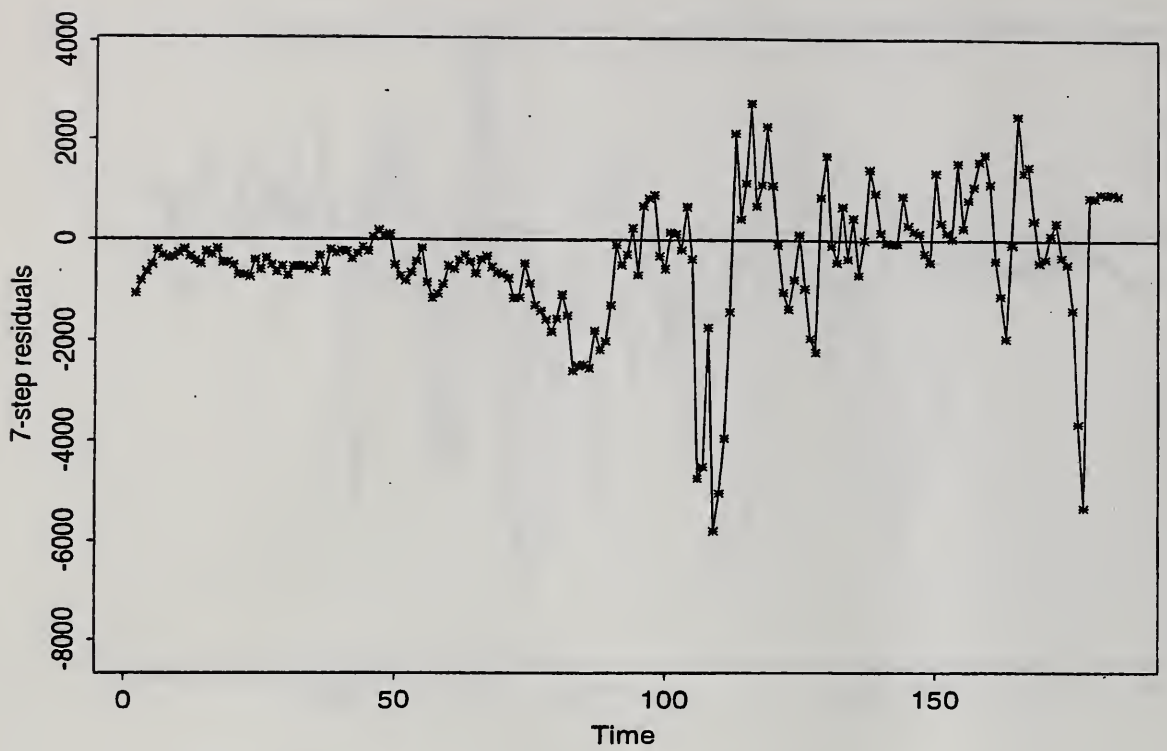

Figure 83: Residuals from Seven-step Predictions - 1995

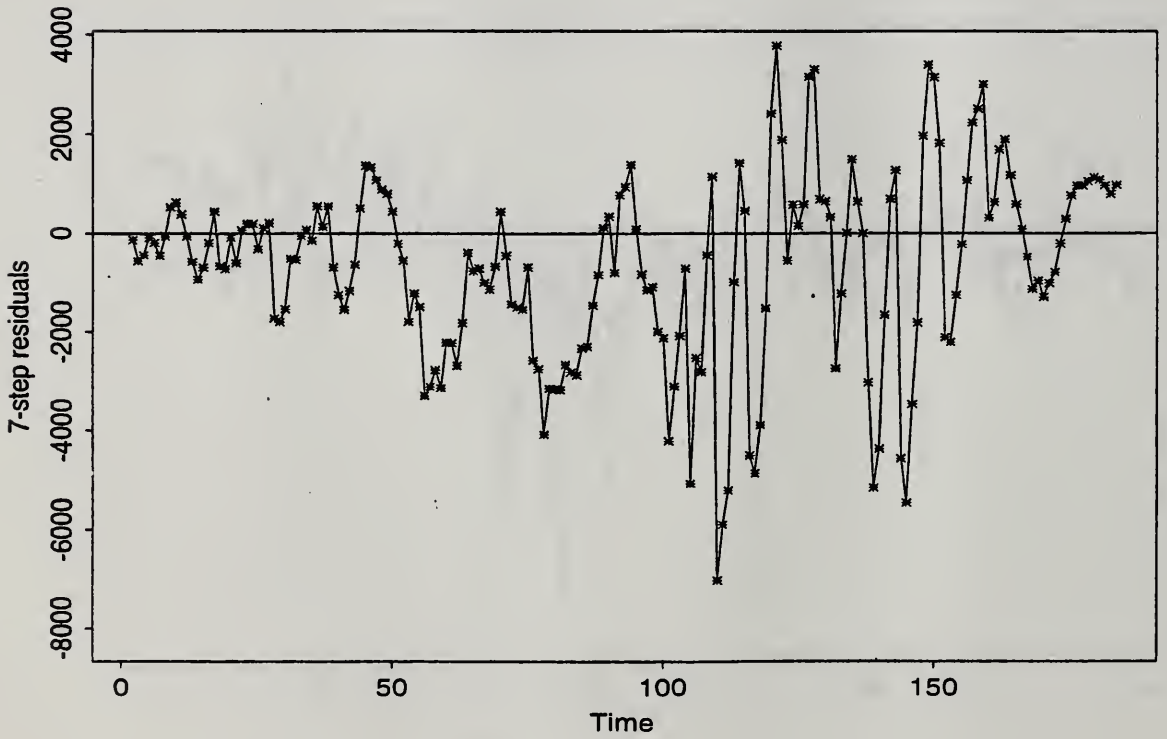

Figure 84: Residuals from Seven-step Predictions - 1996

8 Modeling New Data

Although the focus of this project was on the power data obtained from the Holyoke station, our goal was to develop a model which could be used to make power predictions for other stations in the surrounding area. In this section we examine the predictions under the final model (described in Section 7.2) for two additional power stations, Haxtun and Interstate.

Figures 85-89 give the one-step predictors of the maximum daily power usage for the Haxtun station, Figures 90-94 give the one-step predictors for the Interstate station, Figures 95-99 give the seven-step predictors for the Haxtun station and Figures 100-104 give the seven-step predictors for the Interstate station. In addition, each graph depicts the observed daily maxima for the station under consideration. The time frame used was the same as that for the Holyoke station, namely April 1 to September 30. In order to show the effects of rainfall, solid circles have been plotted indicating a positive amount of precipitation on the previous day. Although the magnitude of the daily maxima was not the same for these stations (the maxima for the Haxtun station are somewhat smaller than for Holyoke whereas the maxima for the Interstate station are roughly double), the model adapted quickly and the predictions were quite good. For the first three years it appears as if missing data at the beginning of the year was filled in with zeros for the Interstate station (see Figures 90-92). Even in this situation, the predictions for the first several days of true data were satisfactory, and the ability of the model to adjust to the sudden change (due both to the missing observations and the magnitude of the observations) is amazing. Within just a few days, the estimates of the polynomial coefficients adapted to adjust for this change.

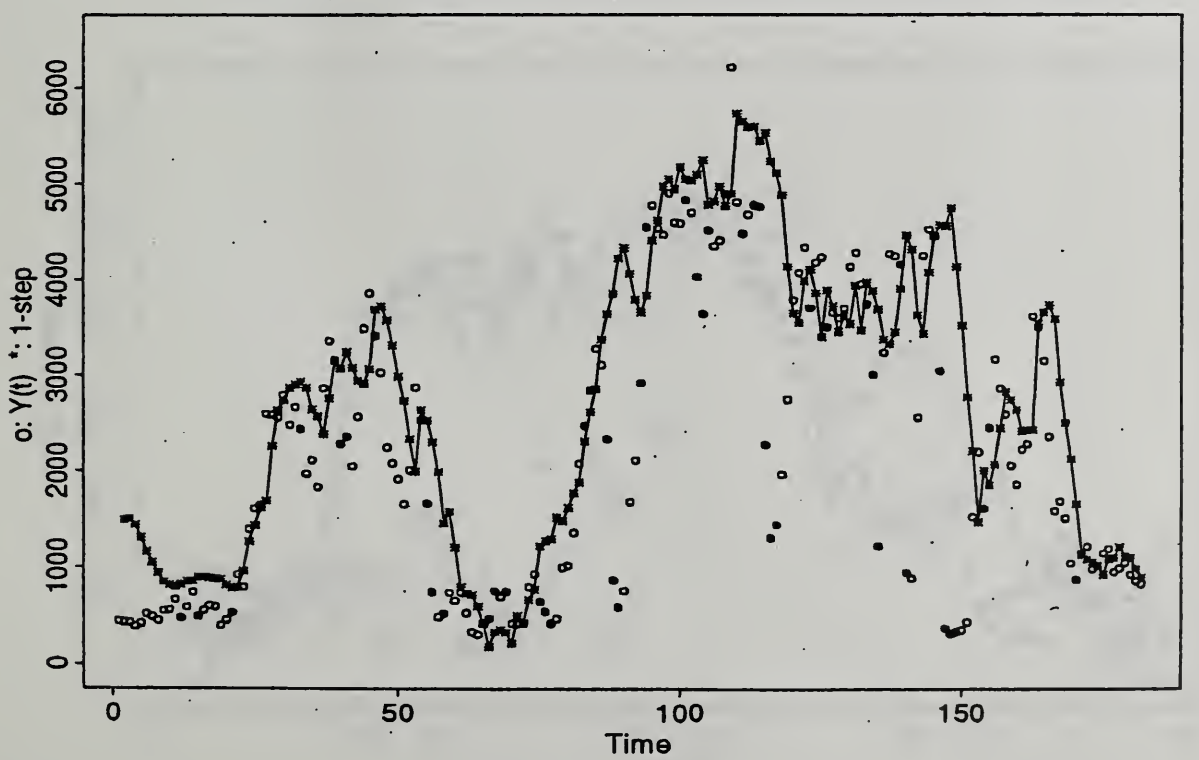

Figure 85: One-step Predictors for the Haxtun Station - 1992

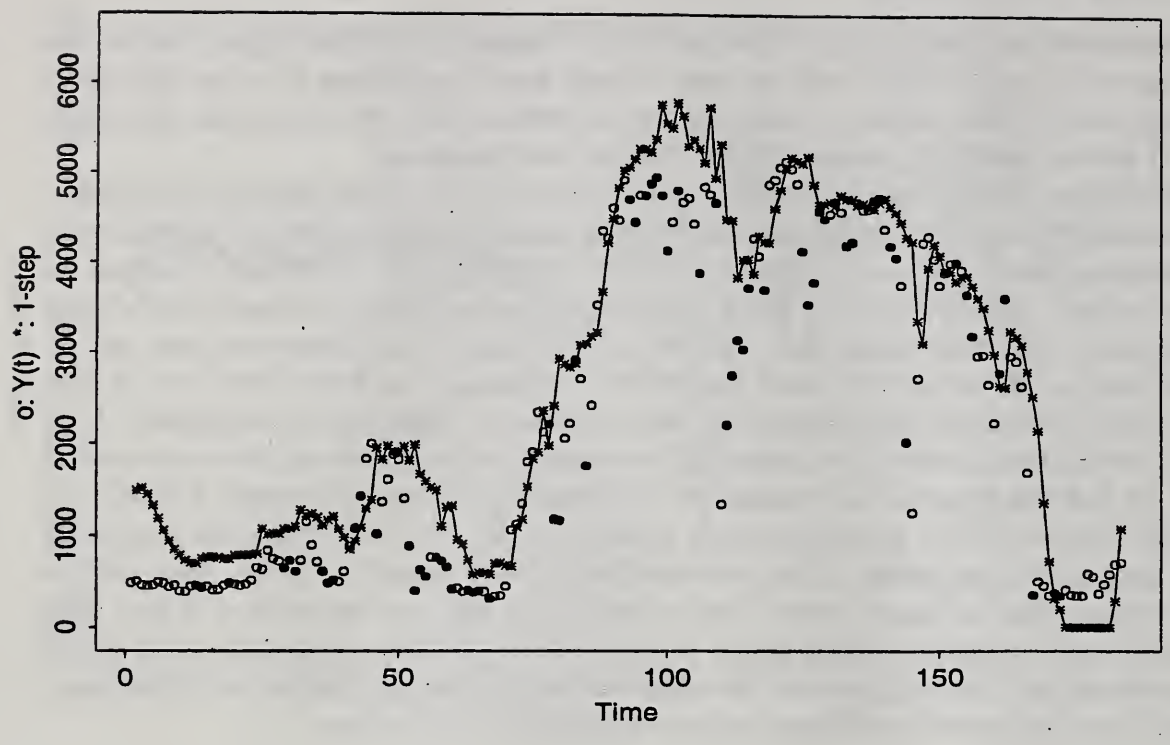

Figure 86: One-step Predictors for the Haxtun Station - 1993

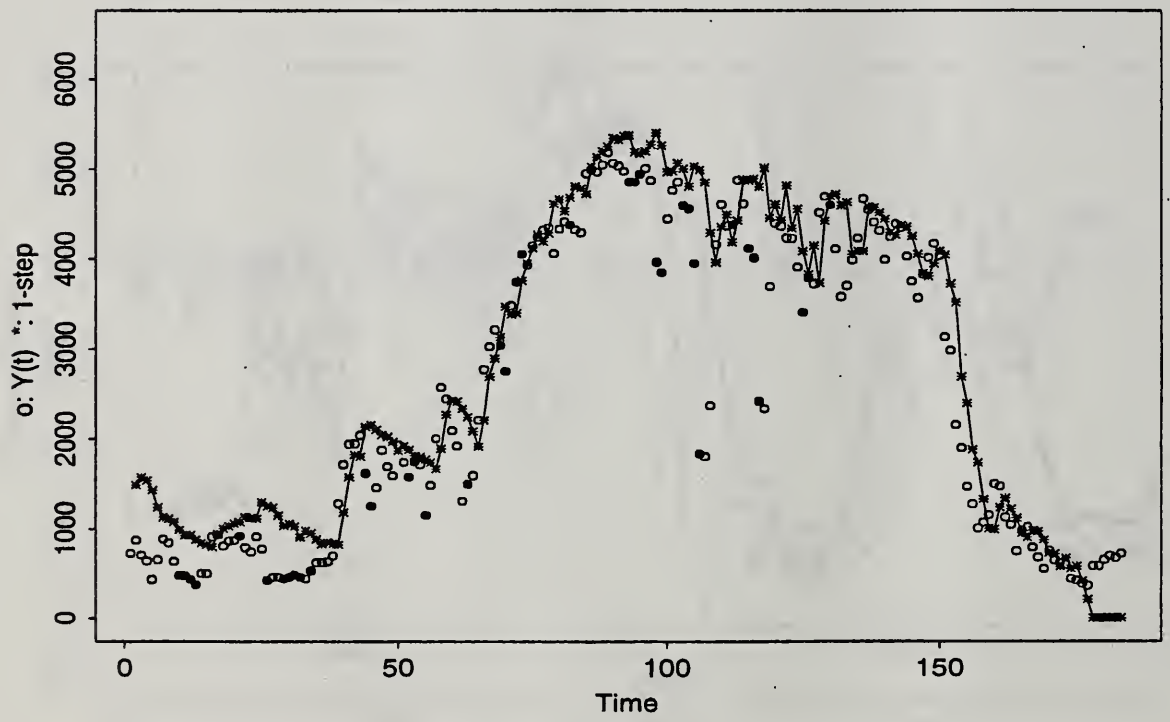

Figure 87: One-step Predictors for the Haxtun Station - 1994

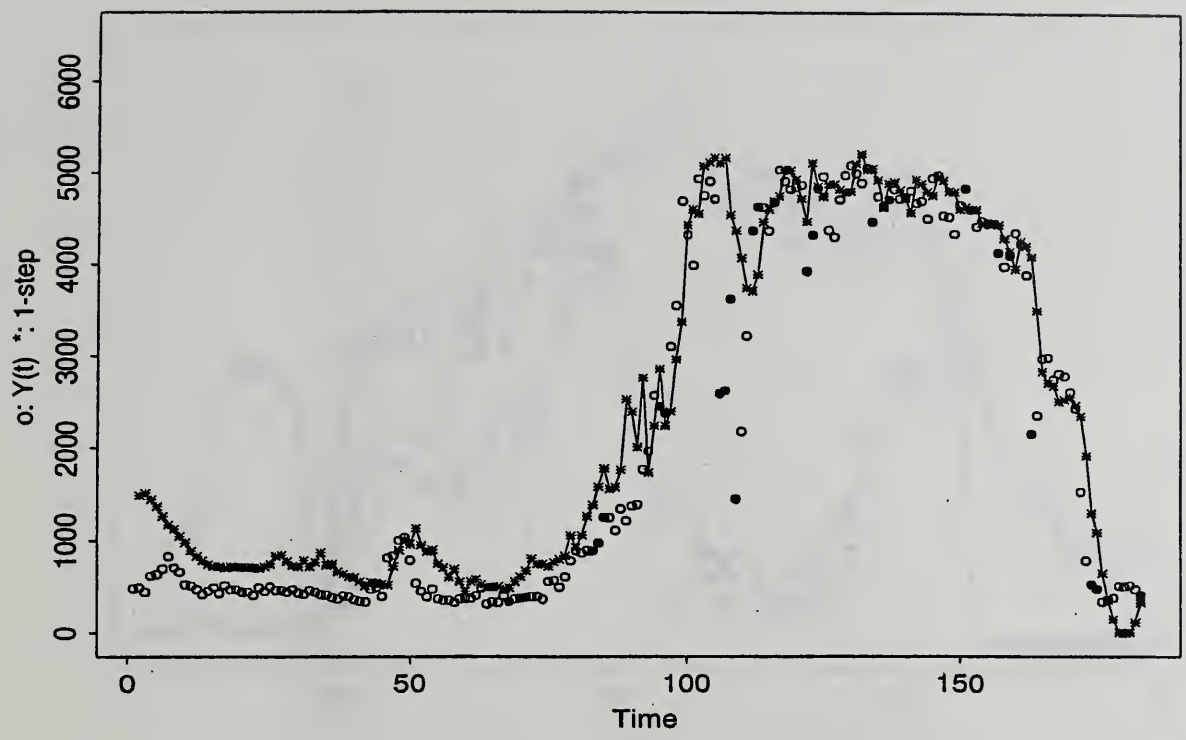

Figure 88: One-step Predictors for the Haxtun Station - 1995

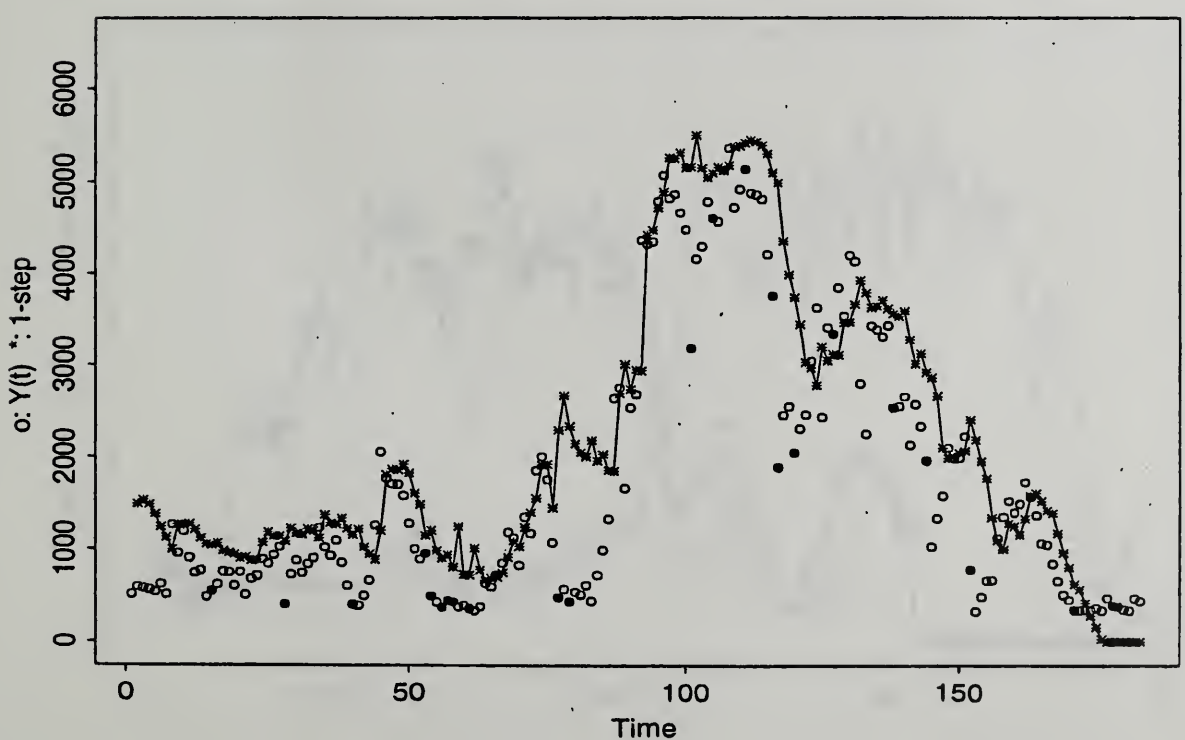

Figure 89: One-step Predictors for the Haxtun Station - 1996

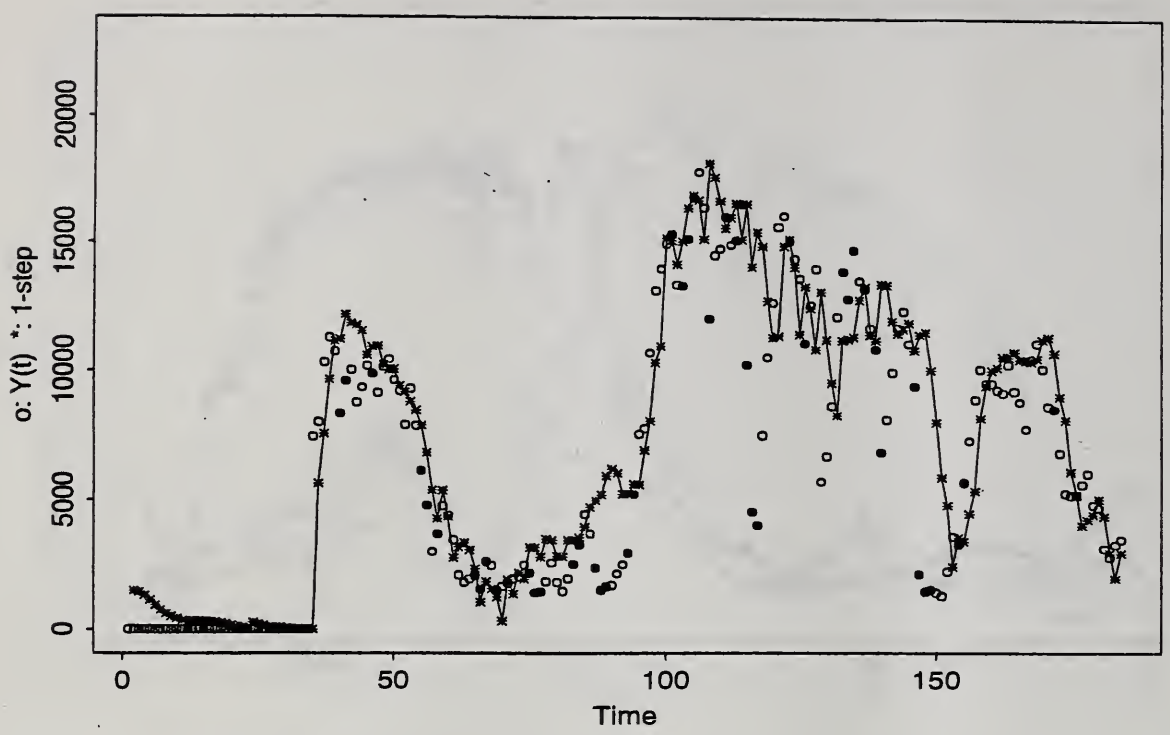

Figure 90: One-step Predictors for the Interstate Station - 1992

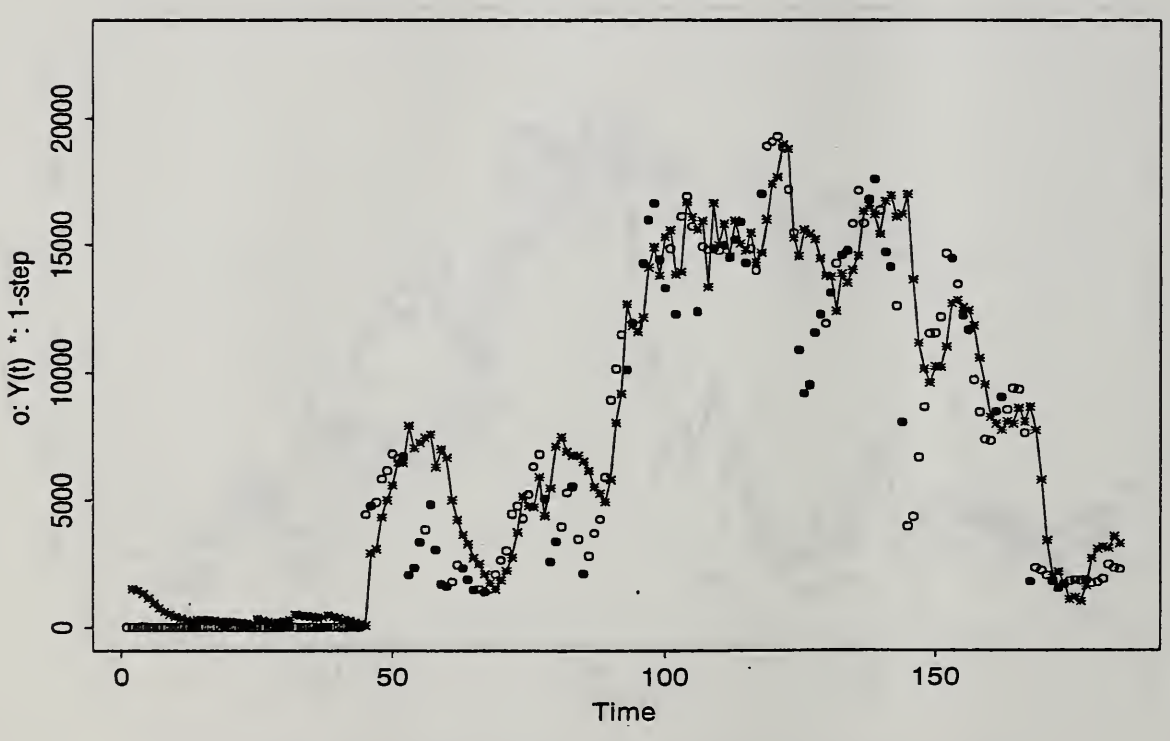

Figure 91: One-step Predictors for the Interstate Station - 1993

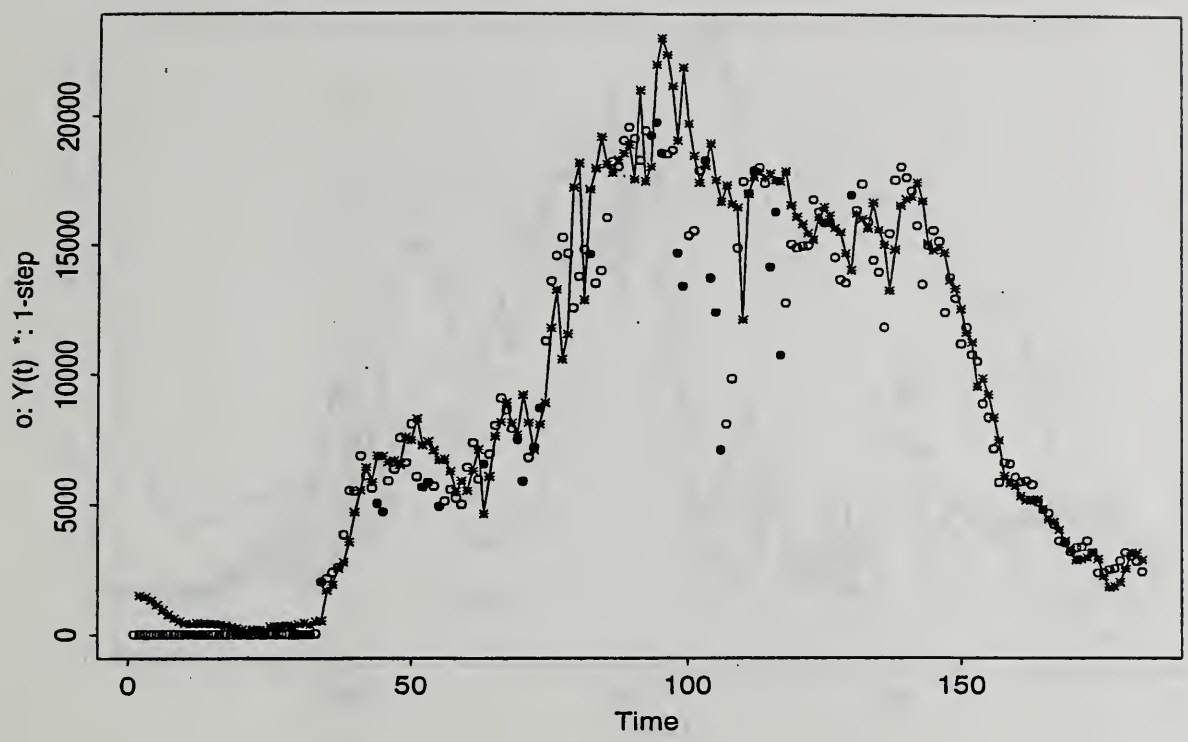

Figure 92: One-step Predictors for the Interstate Station - 1994

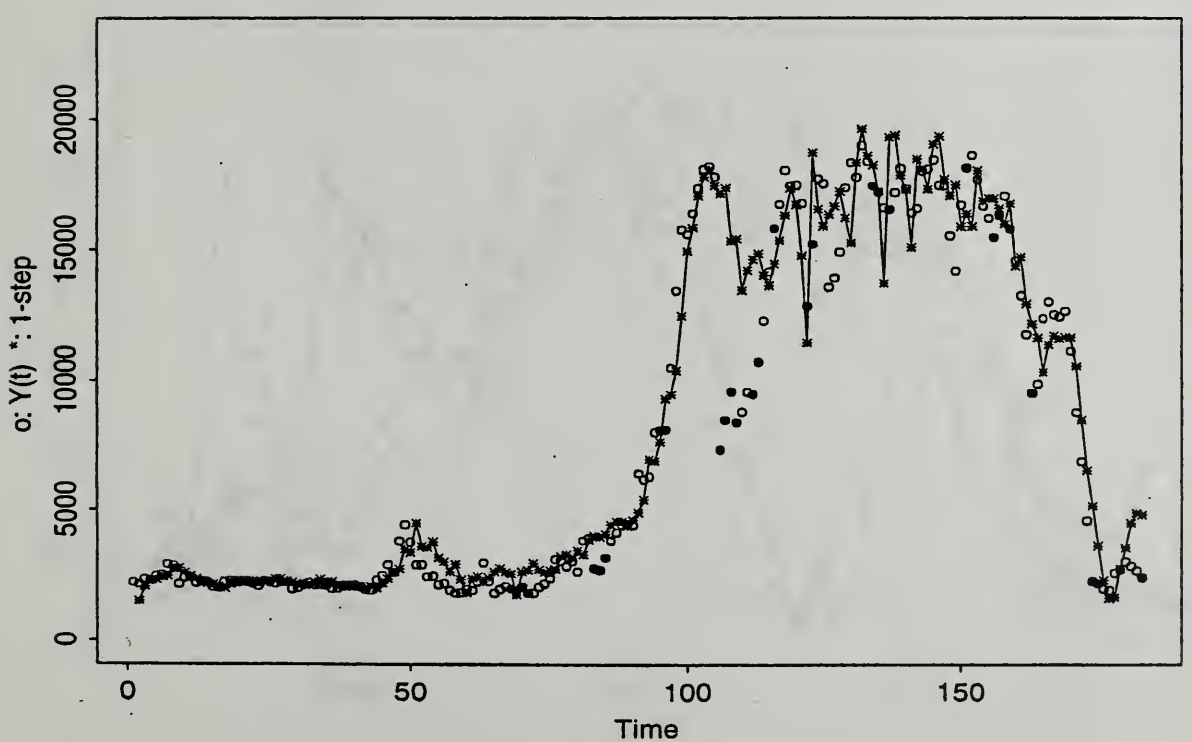

Figure 93: One-step Predictors for the Interstate Station - 1995

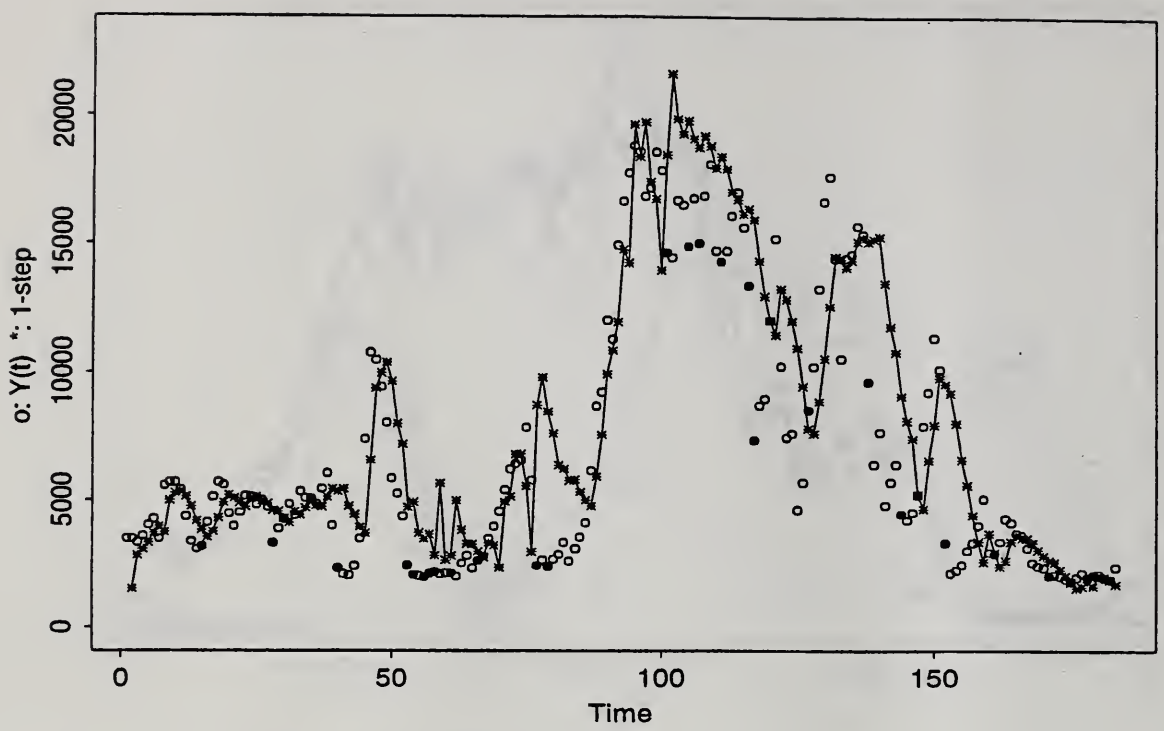

Figure 94: One-step Predictors for the Interstate Station - 1996

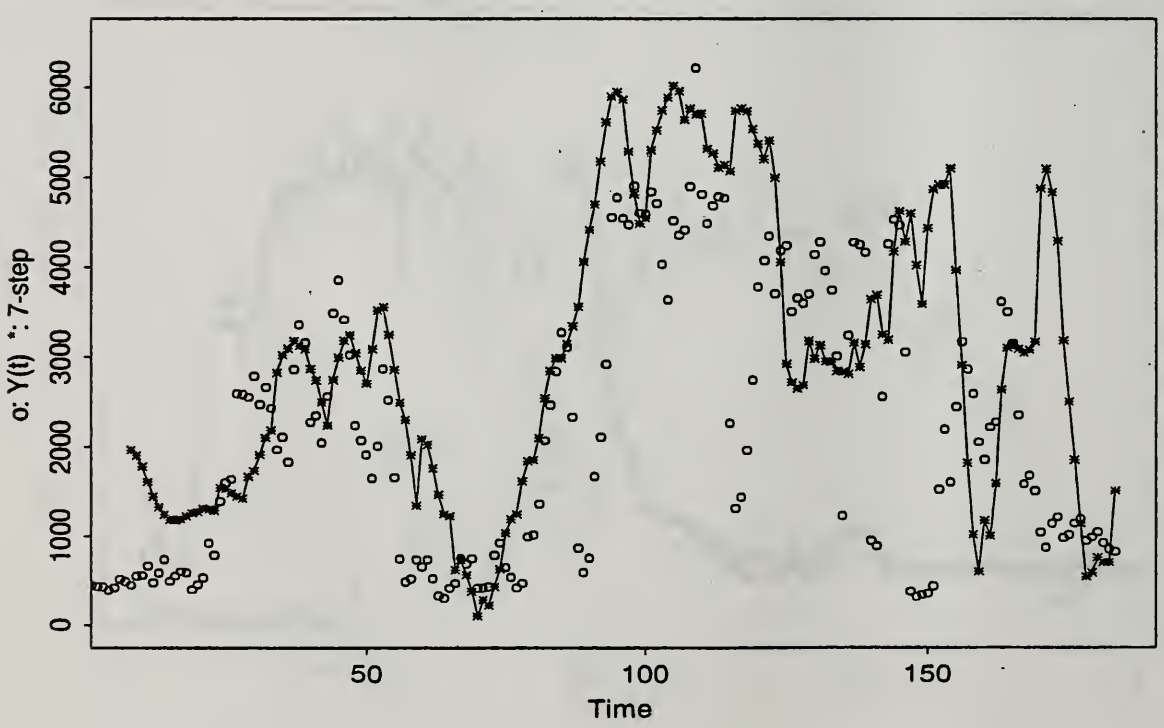

Figure 95: Seven-step Predictors for the Haxtun Station - 1992

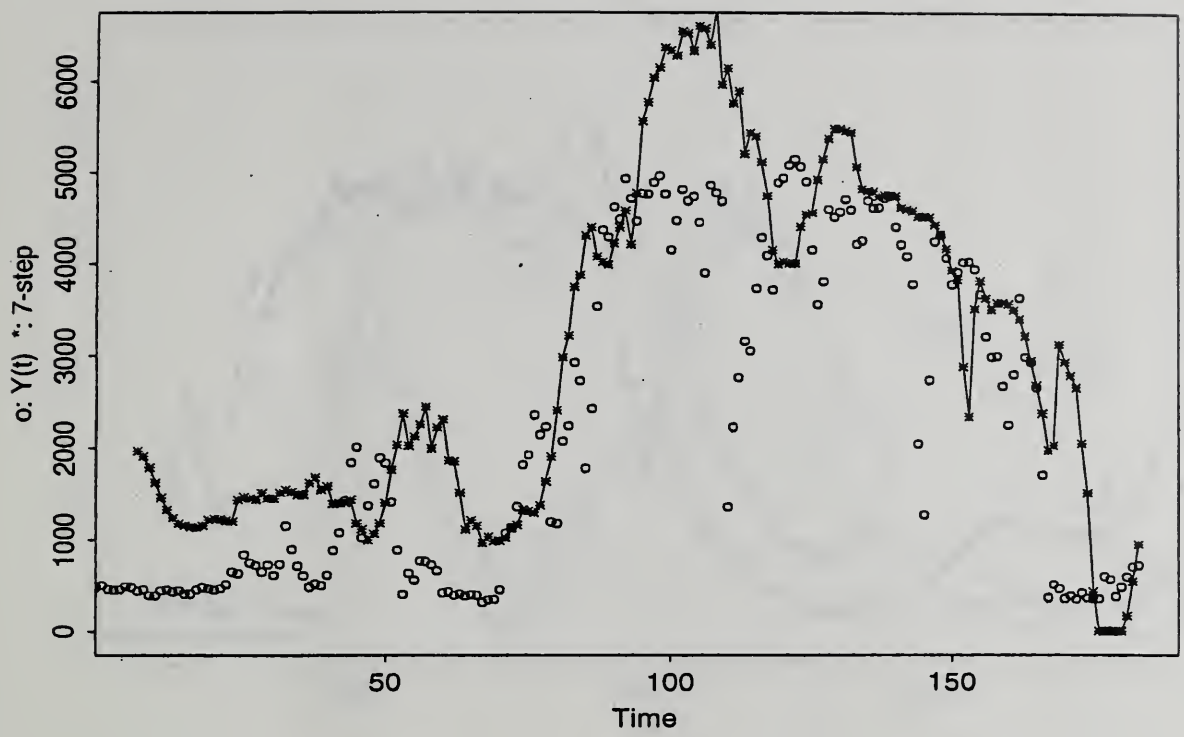

Figure 96: Seven-step Predictors for the Haxtun Station - 1993

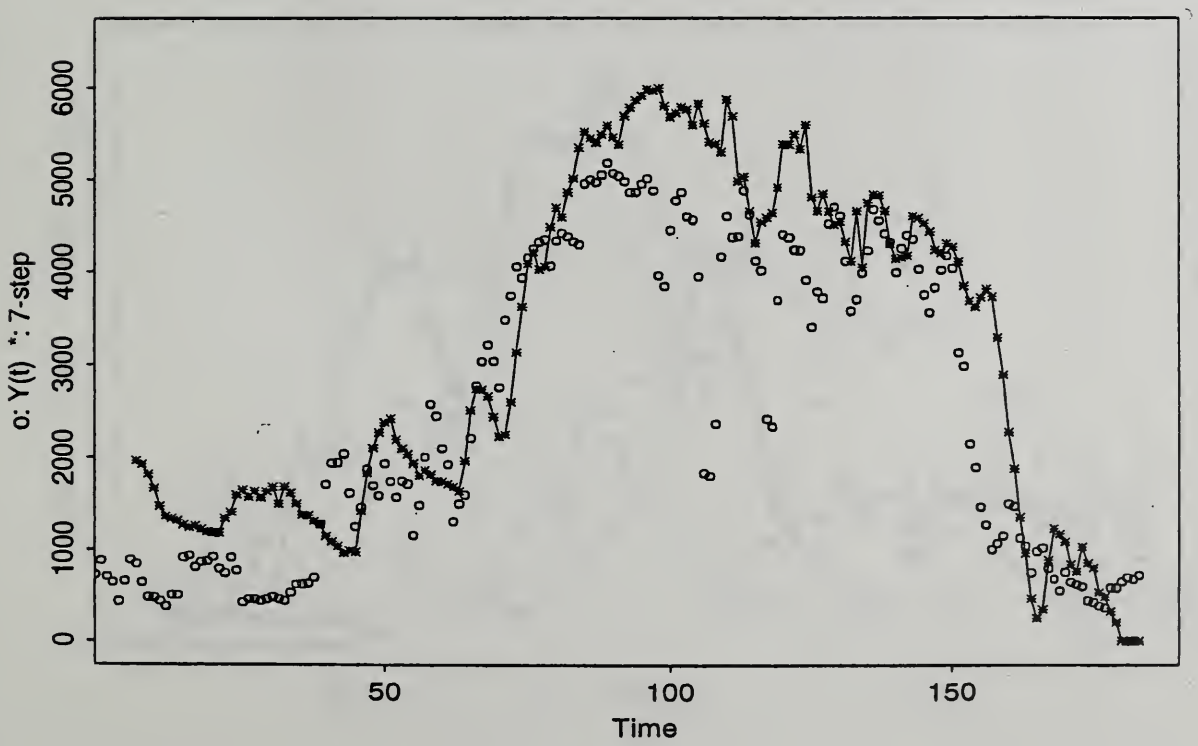

Figure 97: Seven-step Predictors for the Haxtun Station - 1994

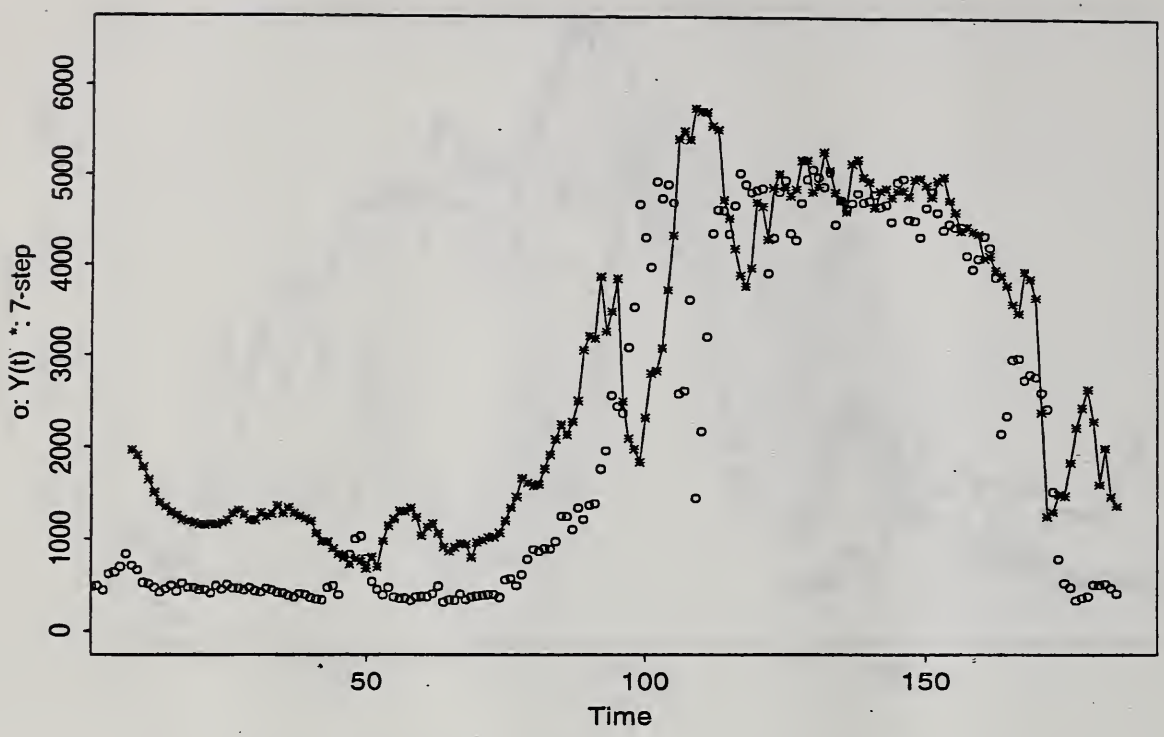

Figure 98: Seven-step Predictors for the Haxtun Station - 1995

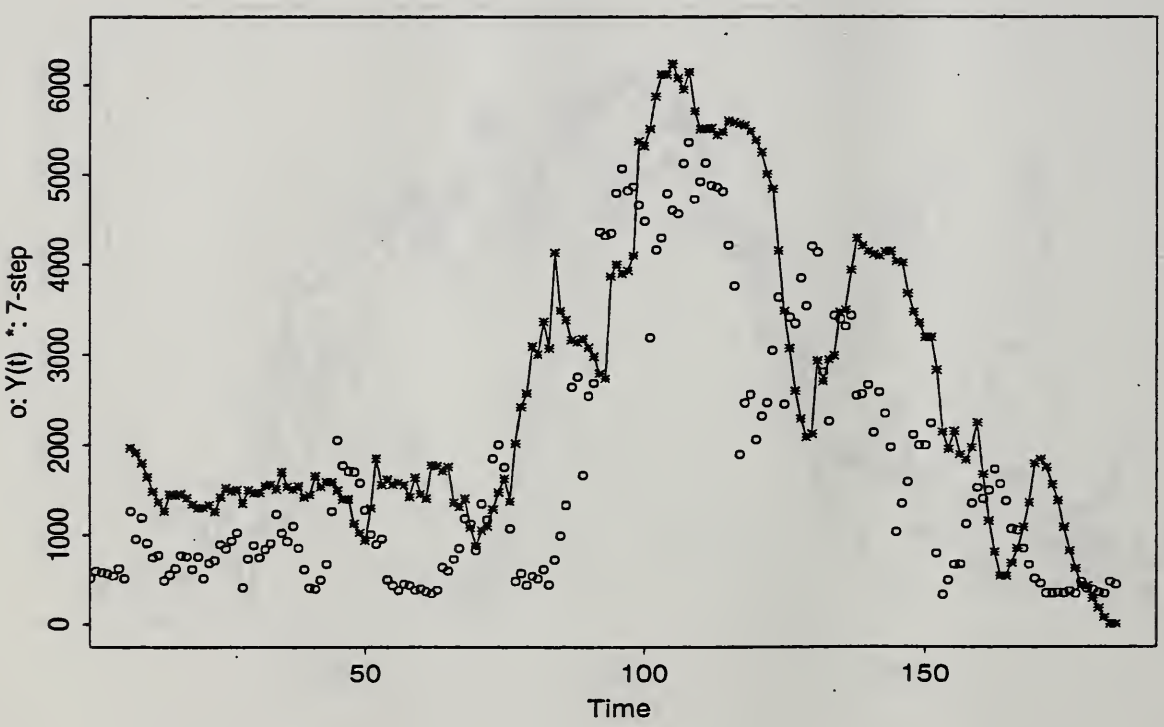

Figure 99: Seven-step Predictors for the Haxtun Station - 1996

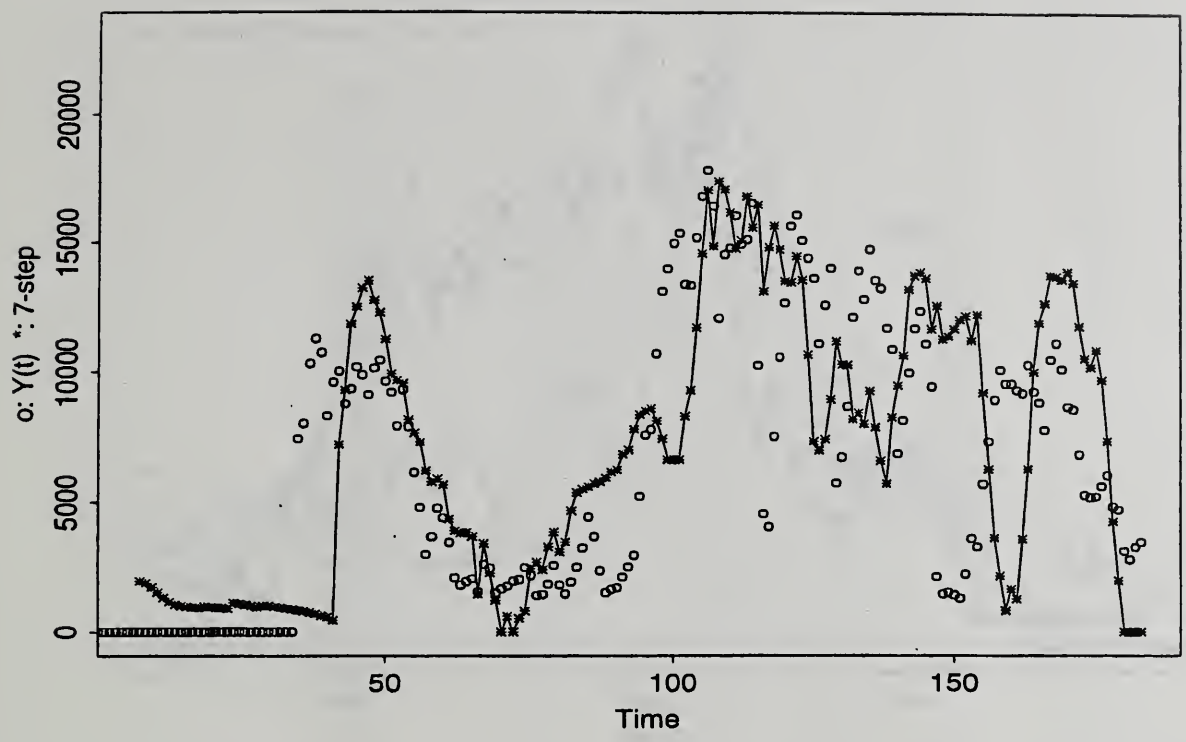

Figure 100: Seven-step Predictors for the Interstate Station - 1992

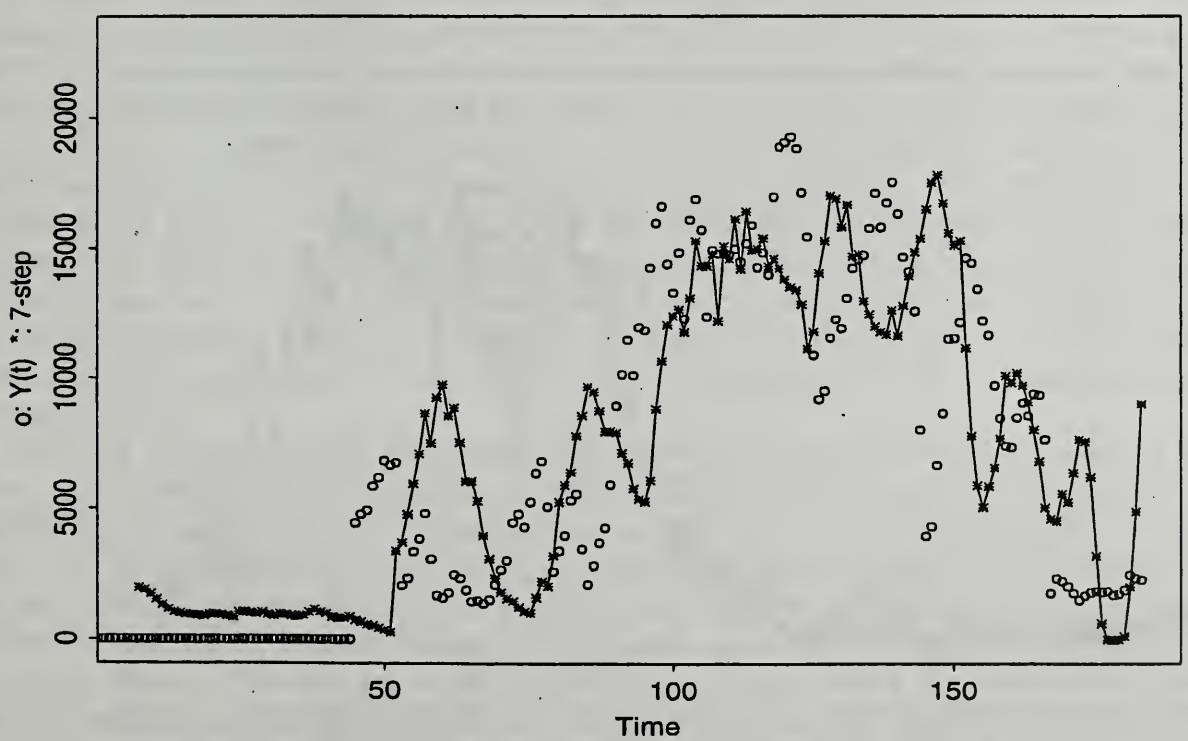

Figure 101: Seven-step Predictors for the Interstate Station - 1993

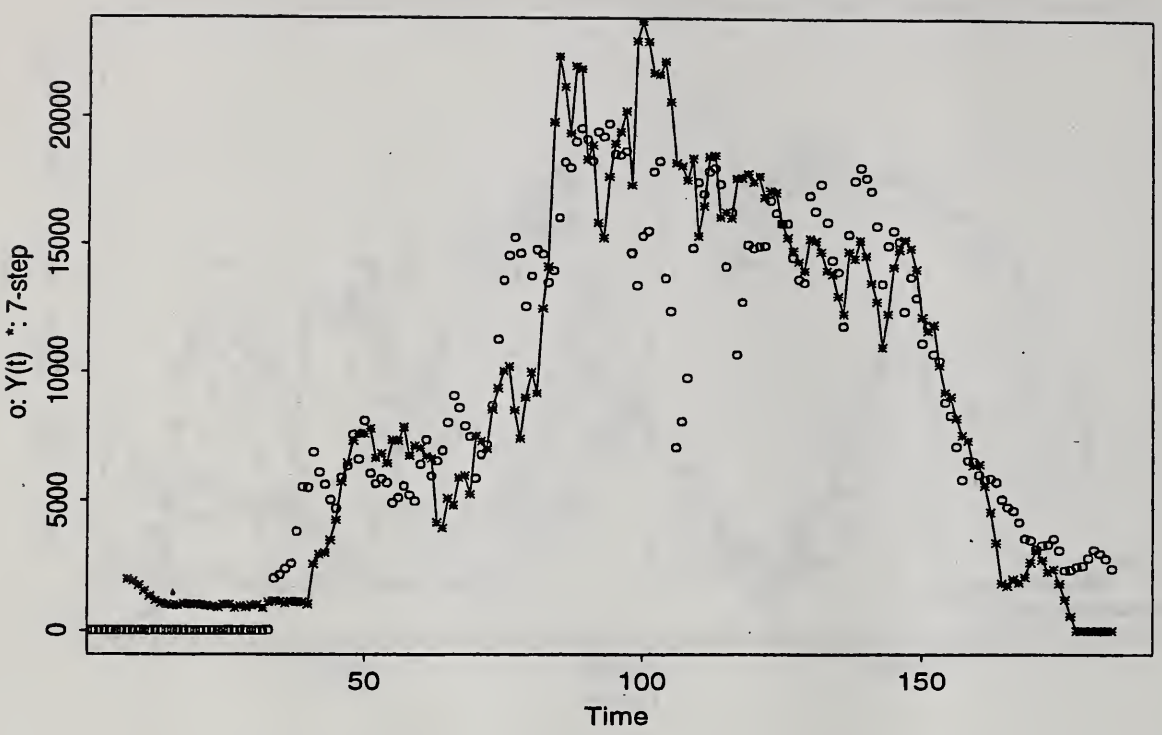

Figure 102: Seven-step Predictors for the Interstate Station - 1994

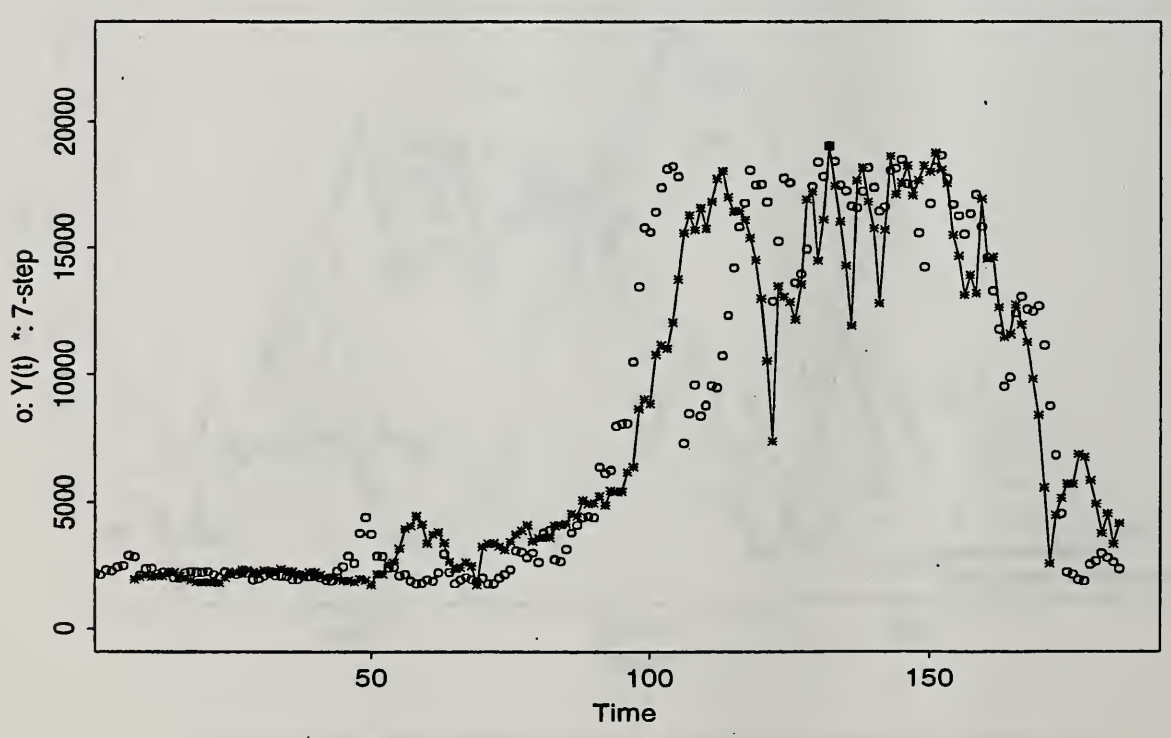

Figure 103: Seven-step Predictors for the Interstate Station - 1995

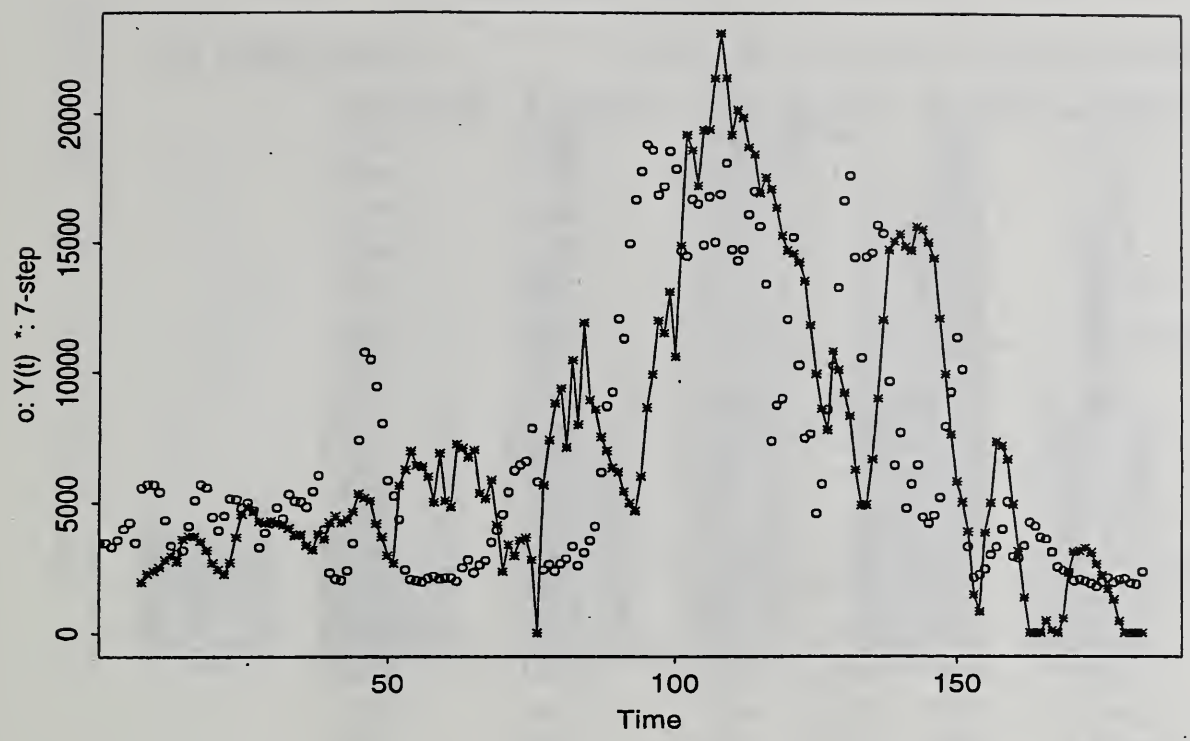

Figure 104: Seven-step Predictors for the Interstate Station - 1996

In order to best utilize the current model, we suggest two prediction strategies. The first, Strategy A, is to be used in the situation when predictions must be made at the beginning of the week for the next seven days. In order to protect against under-predicting the peaks, we suggest that the predictions for the next seven days based on the current day (e.g. the predictions for days $t+1, t+2, \ldots, t+7$ based on day t) all be the same value. This value is to be the maximum of the predictions for days $t+1, \ldots, t+7$ based on day t, day $t-1$ and day $t-2$, as well as the observed maximums for days $t, t-1$ and $t-2$. If it is possible to run the program daily, we recommend Strategy B. During the first month of the growing season, April, Strategy B takes the maximum of the predictions for days $t+1, \ldots, t+7$ based on day t as its prediction for the following day. During the remaining months, May thru September, the prediction for day $t+1$ is found by taking the maximum of the predictions for days $t+1, \ldots, t+4$ based on day t and the observed maximums for days $t, t-1, t-2$.

Since the maximum usage each month was of particular interest, we have focused on the accuracy of its prediction in the table and graphs that follow. Table 8 gives the predictions of each monthly maximum under both modeling strategies, A and B, for the Holyoke data for all five available years. Also included are the 1 and 7-step predictors discussed previously. Comparison of the methods shows that while Strategies A and B are both more accurate than the one and seven-step predictors, Strategy B is to be favored. This is again apparent in comparison of Figures 105-107 to Figures 108-110. The first set of bar graphs (Figures 105-107) show the observed and predicted monthly maximums for all five years for the Holyoke, Haxtun and Interstate stations under Strategy A. The remaining graphs (Figures 108-110) depict the same under Strategy B. Examination of the figures shows that while the predictions under both strategies are quite good, there is a significant amount of improvement in using Strategy B over Strategy A. Considerations on how to further

improve this modeling strategy are given in the following section.

Table 8: Monthly Maximum Predictions for Holyoke

Year	Month	Observed	1-step	7-step	Strategy A	Strategy B
1992	April	3070	2408	1989	2436	2676
	May	4516	3717	2407	2436	3997
	June	3531	2871	3089	4529	3717
	July	9366	9109	8946	9612	9404
	August	8450	6178	4874	7603	7972
	September	4706	4176	3427	4874	5164
1993	April	1146	1240	*	1959	1959
	May	2189	2002	1841	2043	2178
	June	7286	6349	3956	4116	7339
	July	9084	7741	7741	5219	8778
	August	8888	8671	5630	5219	9084
	September	4539	3627	2472	4429	3911
1994	April	1763	1779	1428	1505	2198
	May	2926	2475	2696	2788	2559
	June	9406	9140	10,021	10,284	9350
	July	9538	9069	8430	10,162	10,674
	August	8692	7882	7826	8300	8316
	September	4470	5375	6097	8052	6296
1995	April	1486	1611	*	1959	1709
	May	1613	1399	1118	1222	1499
	June	2310	2255	3590	3933	2984
	July	9700	8842	5284	6282	8922
	August	9308	9178	9866	10,107	9274
	September	7770	7512	7896	8842	8398
1996	April	3036	2780	2135	2139	2836
	May	4320	3286	2572	2663	3629
	June	5414	4977	3750	4382	5533
	July	9740	9412	10,746	10,719	9534
	August	8876	7826	6897	7585	8518
	September	3012	1064	3096	4514	2517

* the monthly maximum occured before the 7th day

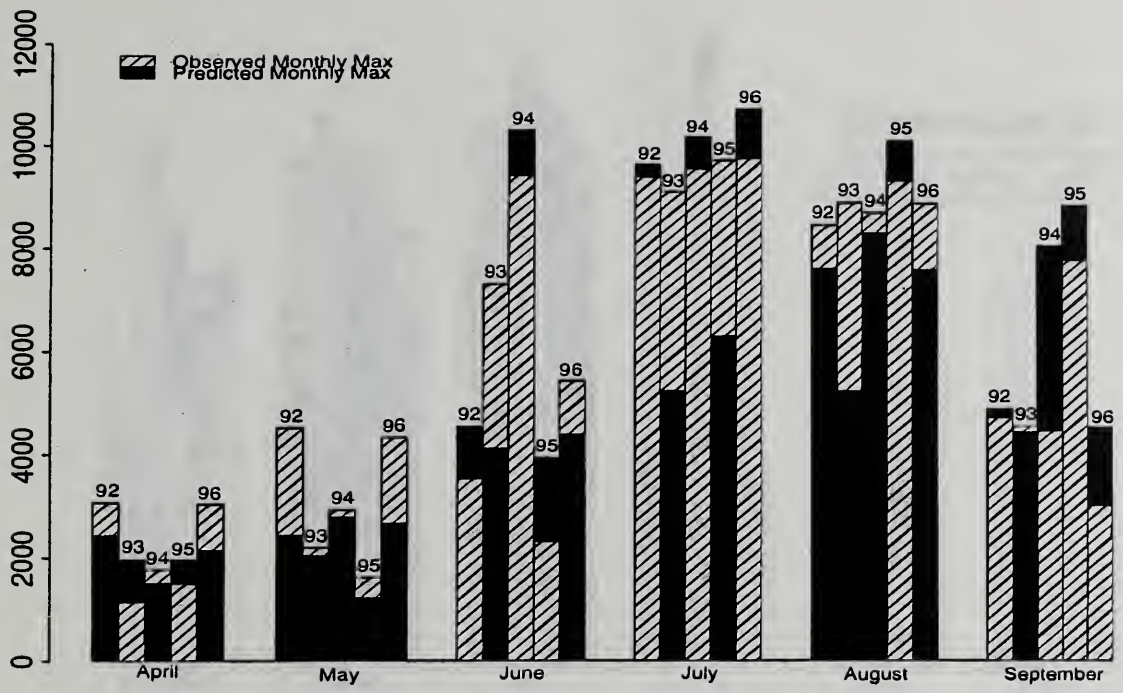

Figure 105: Monthly Maximums for Holyoke Under Strategy A

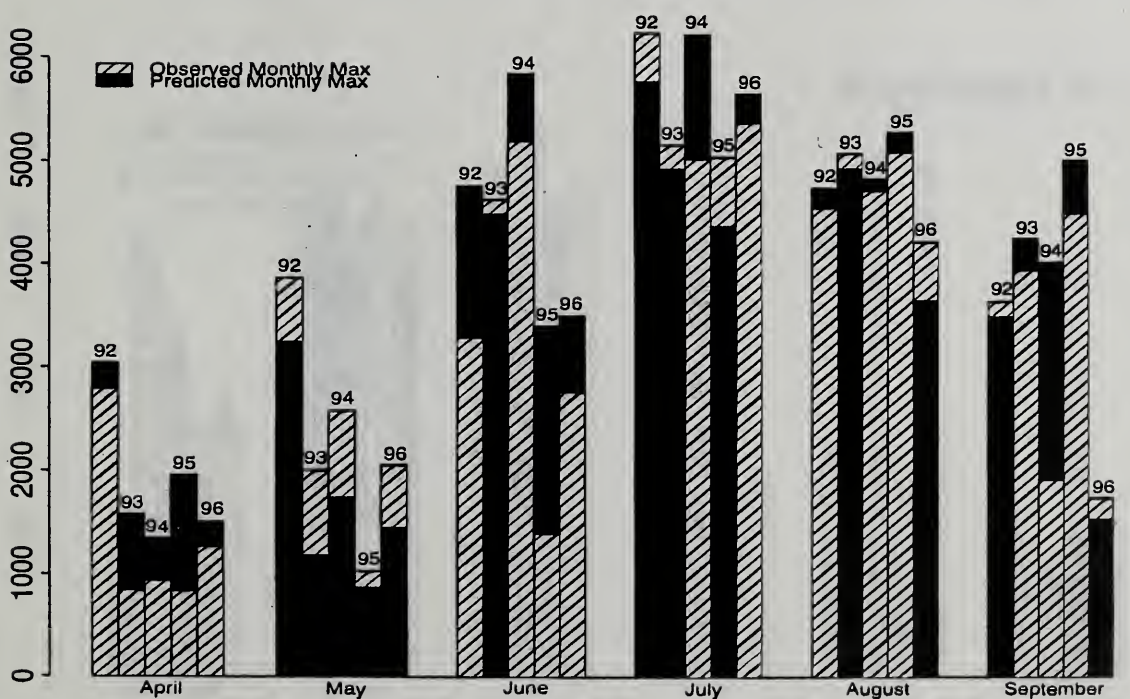

Figure 106: Monthly Maximums for Haxtun Under Strategy A

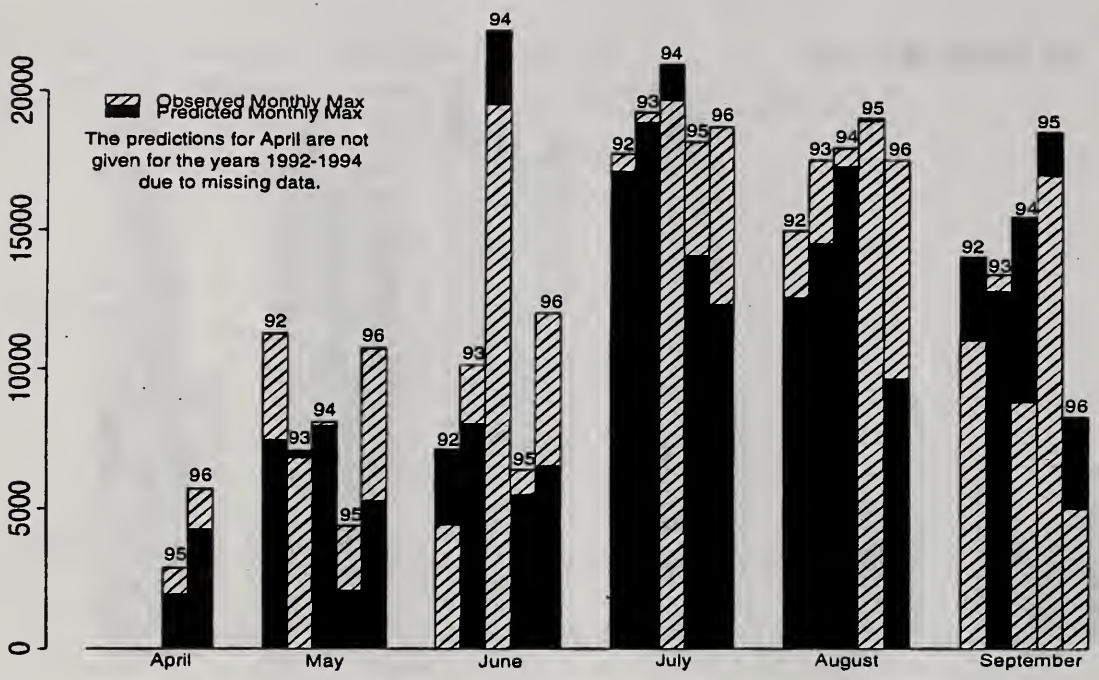

Figure 107: Monthly Maximums for Interstate Under Strategy A

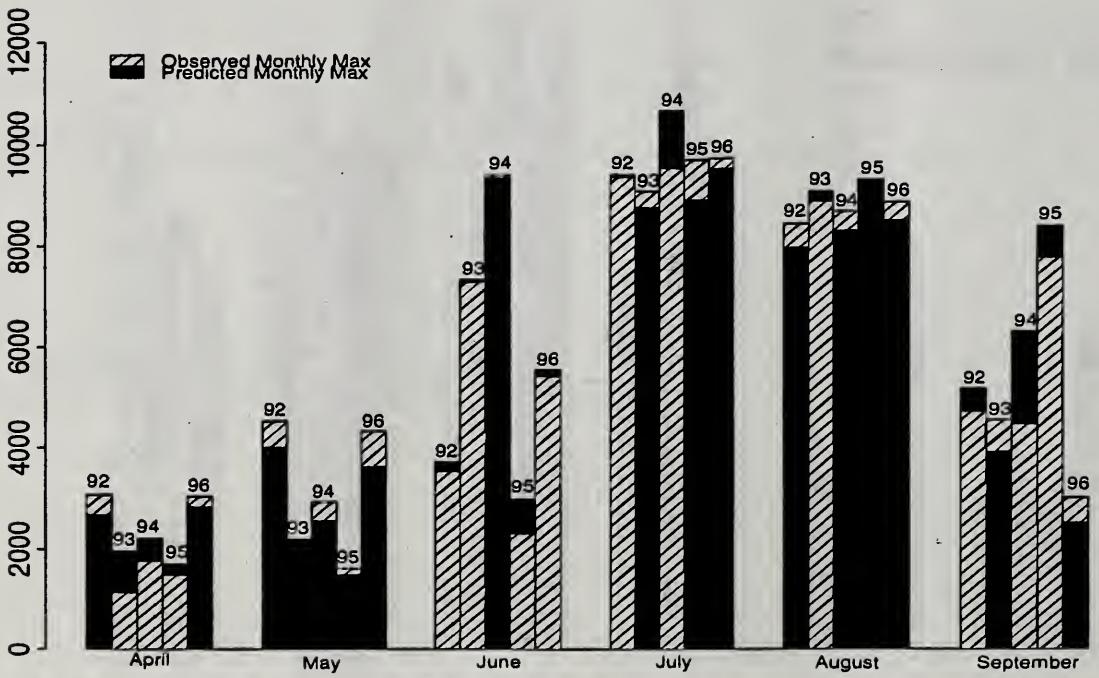

Figure 108: Monthly Maximums for Holyoke Under Strategy B

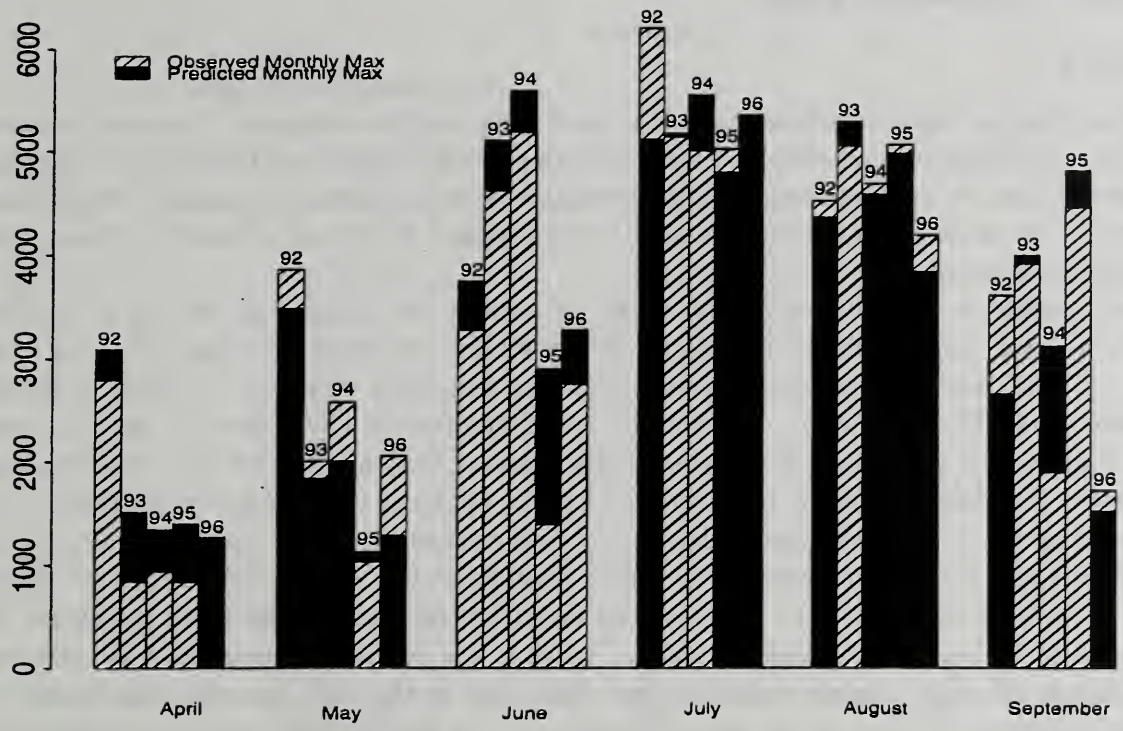

Figure 109: Monthly Maximums for Haxtun Under Strategy B

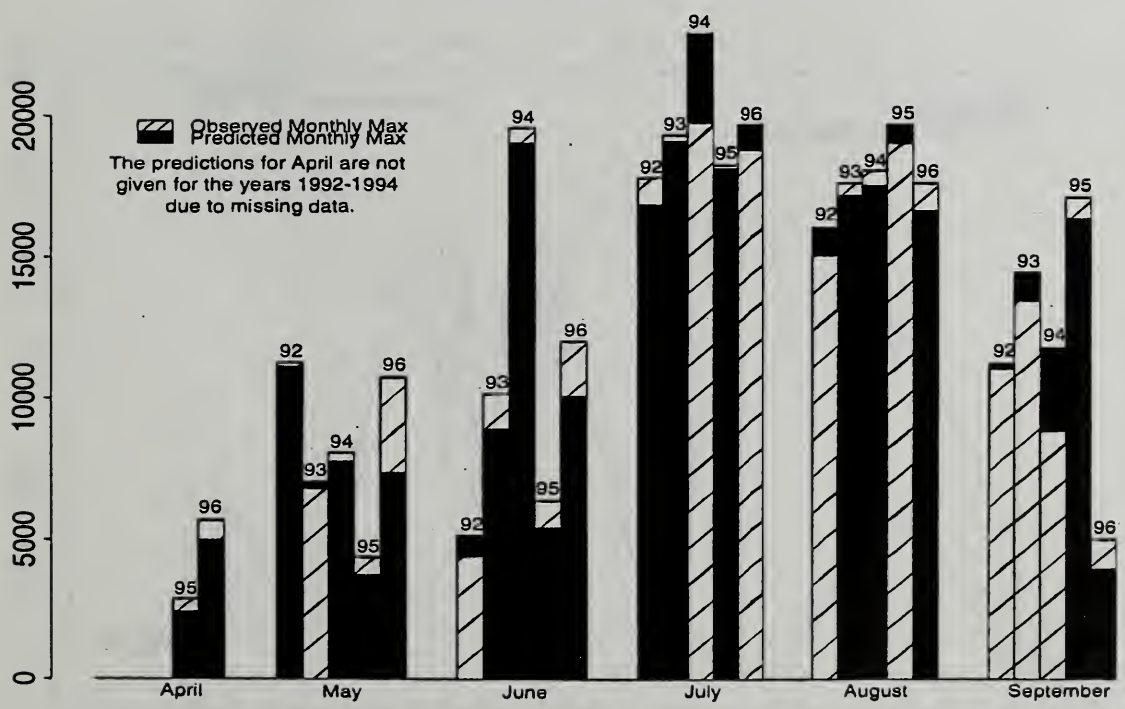

Figure 110: Monthly Maximums for Interstate Under Strategy B

9 Further Considerations

9.1 Wavelets

One of the directions we have considered taking in predicting the data concerns the use of wavelet functions in lieu of Legendre polynomials. Wavelets were considered because, theoretically, a small number of wavelets can be used to closely model the data even when spikes are present. Polynomial functions tend to be much smoother than wavelets which makes it difficult to model erratic data in detail without the use of many functions.

A wavelets module is available for S-Plus which we used in our preliminary fitting of wavelet functions to the data. *Applied Wavelet Analysis with S-PLUS* by Bruce and Gao ([3]) was used as a reference to this module. The module contains several functions which can be used in fitting wavelets to the data. We tried many of these and found the function *best.basis* to be the most satisfactory. Figures 111-115 depict the original series (dashed line) along with the reconstructed series (solid line). The reconstruced series was found by first applying the *best.basis* function to the original series using the *d*8 wavelet, abstracting the top ten wavelet functions (those containing the most information) and then reconstructing the series from these ten wavelet functions (see Bruce and Gao [3] for further information). However, we came across some of the same problems we had encountered in fitting the polynomial model. For one, the robust fit discounted all extreme values, large and small. Also, the wavelet functions which best fit the data, among those functions considered, differed from one year to the next. At this point, we decided that there was insufficient time to explore these problems further. However, since the preliminary plots were encouraging, wavelet functions are mentioned here as a possible route to consider in further studies involving similar data sets.

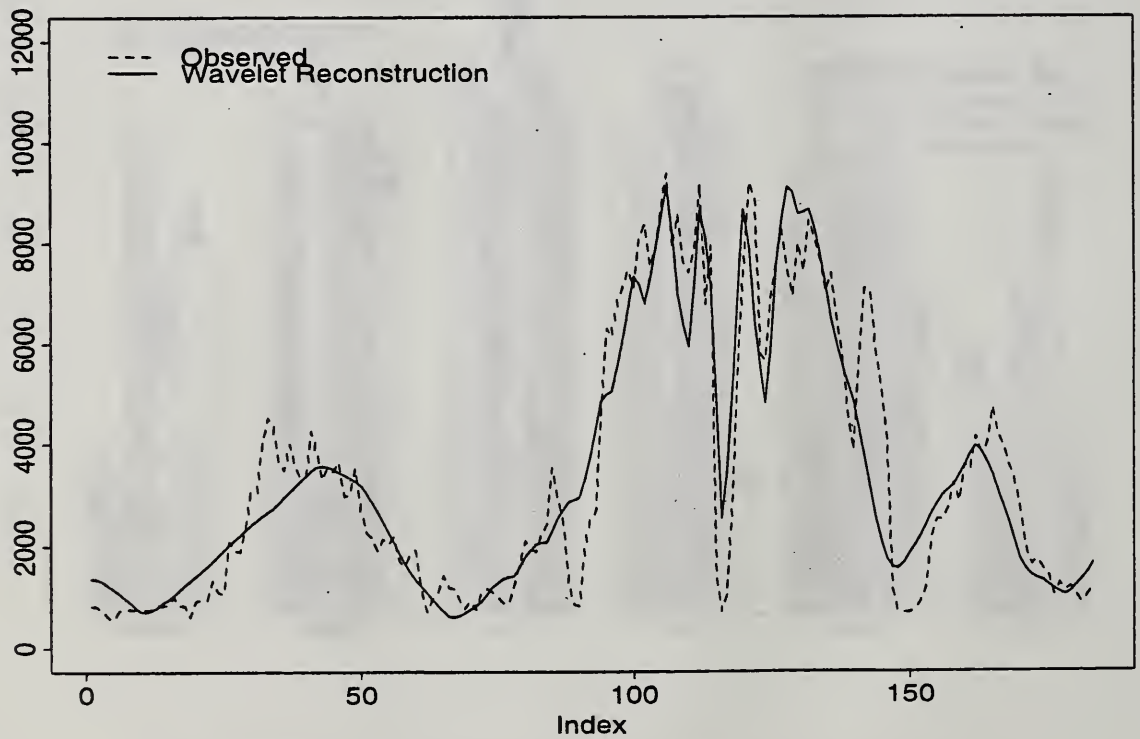

Figure 111: Wavelet Reconstruction - 1992

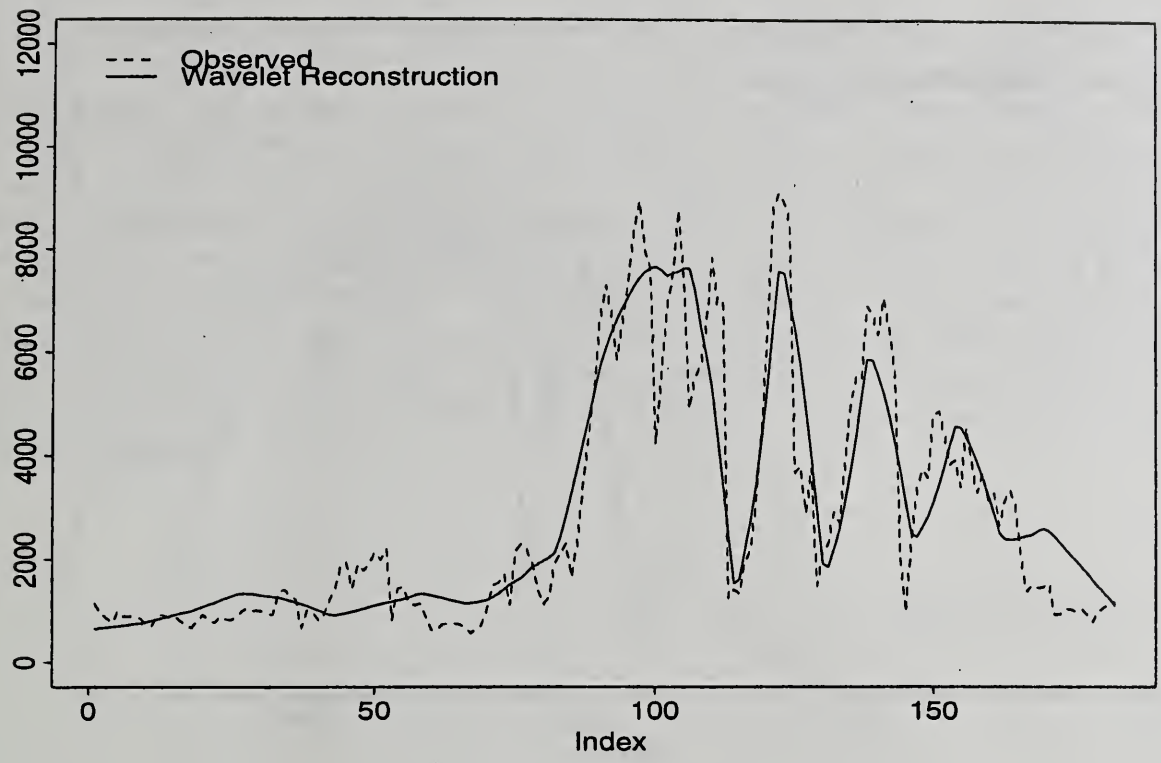

Figure 112: Wavelet Reconstruction - 1993

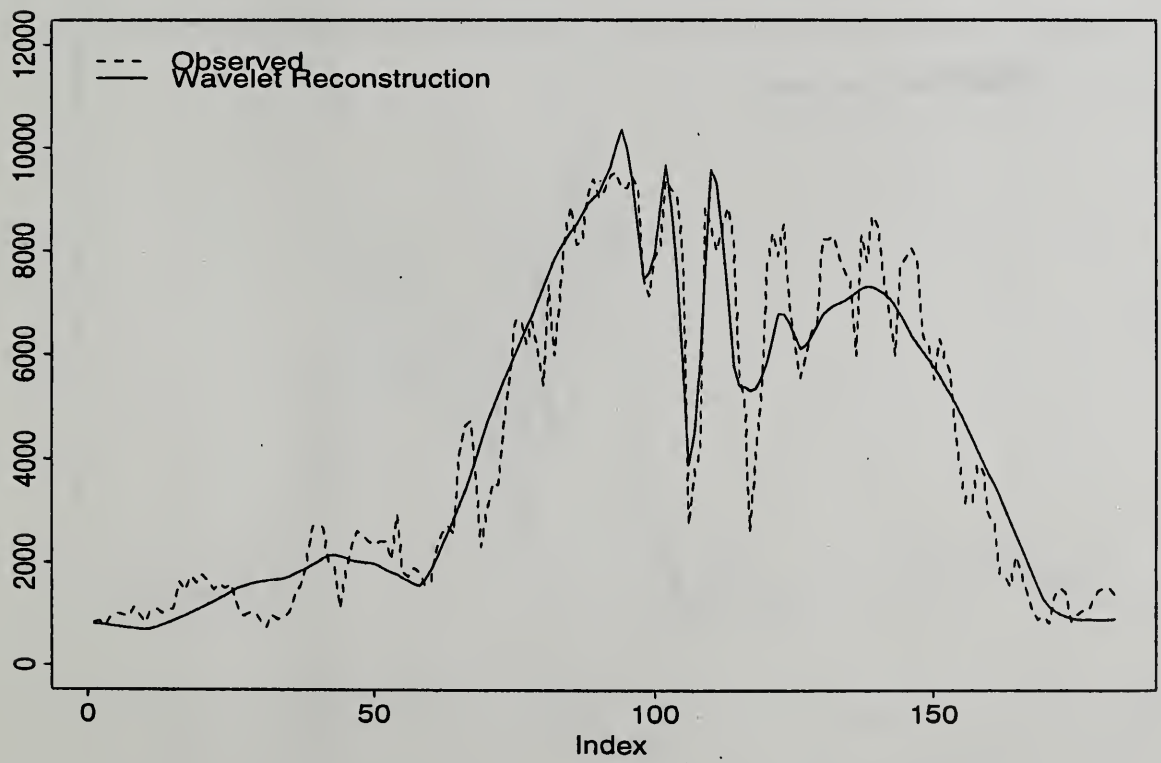

Figure 113: Wavelet Reconstruction - 1994

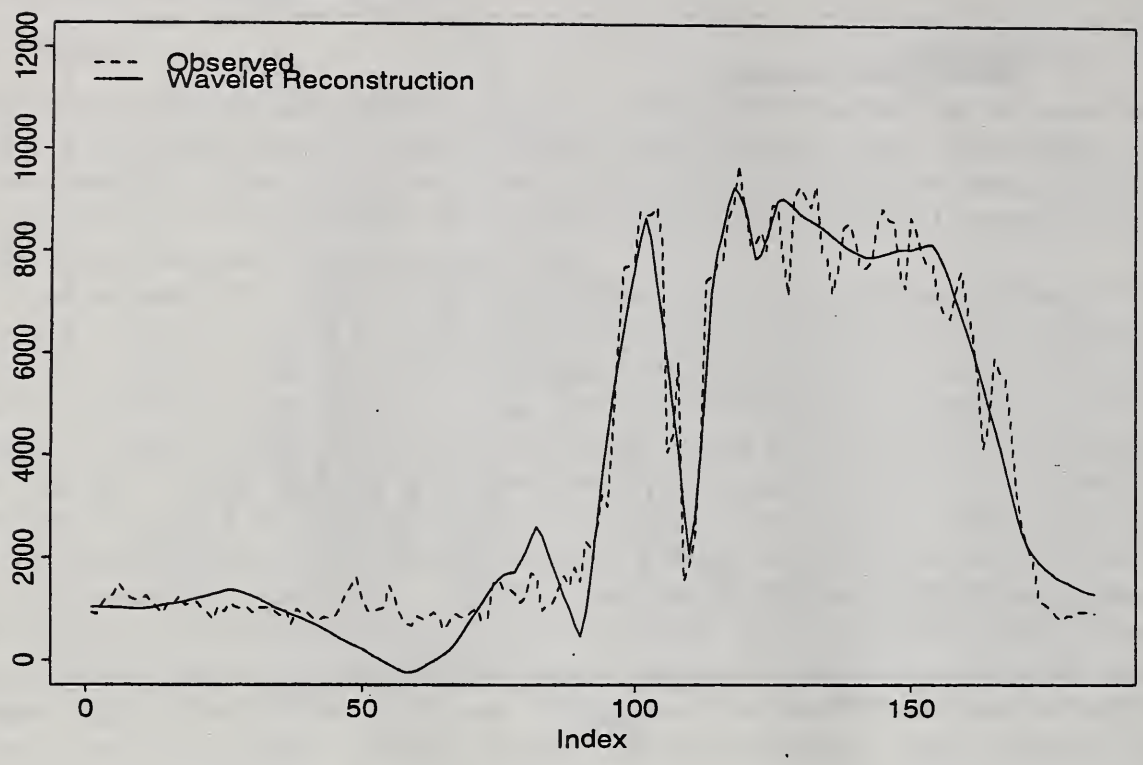

Figure 114: Wavelet Reconstruction - 1995

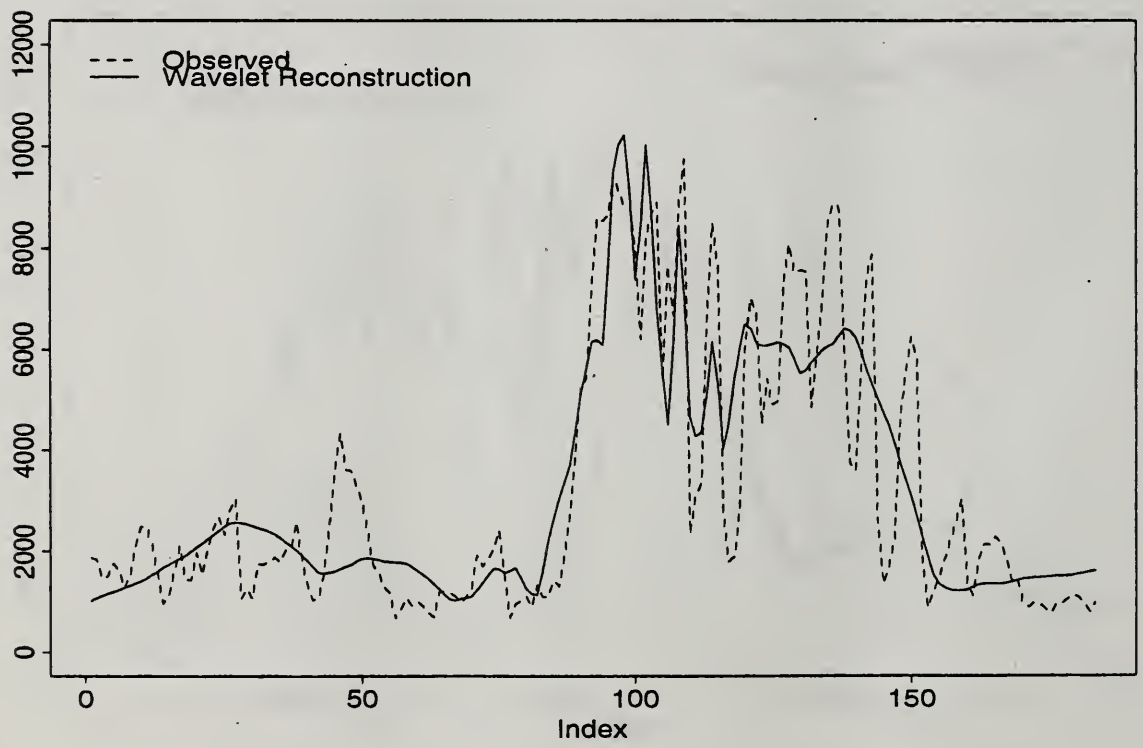

Figure 115: Wavelet Reconstruction - 1996

9.2 Spatial Analysis

Another direction we have not yet explored is the use of a spatial factor in our model. At this point, only the weather variables from one of the "nearby" stations have been integrated into our modeling scheme. For many of the power stations, there are other weather stations which are closer than the Holyoke weather station. The weather data collected from these stations could be combined to provide more information resulting in better predictions. Additionally, the distance between the power station and the Holyoke weather station is much further for some stations than for others. This could be taken into account by allowing weather variables to have a diminishing affect based on the distance between the weather and power stations. It is also likely that the daily maxima for power stations that are close to one another are more strongly related than the maxima for stations that are far apart. Again, this potential relationship could be included in our model by way of a spatial component. The idea of a spatial element is mentioned here as a consideration for further analyses.

References

[1] *The AGEND Irrigation Model User's Guide - The Long Term Model* (AGFORE) (Version 0.1RG), prepared by Quantum Consulting Inc. 2030 Addison Street, Berkeley, CA 94704 tel: (415) 540-7200 for the Electric Power Research Institute, 3412 Hillview Avenue, Palo Alto, CA 94303; Research Project 2340-2., Final User's Guide, October 1990.

[2] Peter Brockwell and Richard Davis. *Introduction to Time Series and Forecasting,* Springer-Verlag, New York, 1996.

[3] Andrew Bruce and Hong-Ye Gao. *Applied Wavelet Analysis with S-PLUS,* Springer-Verlag, New York, 1996.

[4] Buchleiter, Gerald W., 1986. *Optimal Pump Selection for Large Scale Irrigation Systems*, Ph.D. Dissertation Department of Agricultural and Chemical Engineering, Colorado State University, Fort Collins, Colorado, Fall 1986, 87 p.

[5] Buchleiter, Gerald W., 1979. *Water Management Alternatives for Controlling Peak Electrical Demands*, Ph.D. Dissertation, Department of Agricultural Engineering, Colorado State University, Fort Collins, Colorado, Fall 1979, 82 p.

[6] Tomáš Cipra and Rosario Romera. *Robust Kalman Filter and Its Application in Time Series Analysis,* Kybernetika, Volume 27 (1991), Number 6.

[7] Electric Power Software Center (1990) Installation requirements for AGEND - Short Term Model RP-2340-2 IBM PC Version 1.0 PRODUCTION. Electric Power Software Center, 1930 Hi Line Drive, Dallas, Texas 75207, tel: (214) 655-8883 fax: (214) 655-8772, August 1990, 6 p.

[8] P.J. Huber. *Robust Statistics,* J. Wiley, New York, 1981.

[9] E.L. Lehmann. *Theory of Point Estimation,* Springer-Verlag, New York, 1983.

[10] Uhlaner, R.T. and Cason, T.N. (1990). *The Agend Irrigation Model User's Guide - The Short Term Model*, prepared by Quantum Consulting Inc. 2030 Addison Street, Berkeley, CA 94704 tel: (415) 540-7200 for the Electric Power Research Institute, 3412 Hillview Avenue, Palo Alto, CA 94303; Research Project 2340-2., Final User's Guide, January 1990.

[11] Young, Bradley J., 1979. *Controlling Peak Electrical Demands by Irrigation Management*, M.S. Engineering, Department of Agricultural Engineering, Colorado State University, Fort Collins, Colorado, Spring 1979, 82 p.